GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: December 10, 2002, 09:57:25 ; Search time 12 Seconds
(without alignments)
610.440 Million cell updates/sec

Title: US-09-838-028-2
Perfect score: 2318
Sequence: 1 MESSPIPQSSGNSSTLGRVP.....SDIIMSDSYLRPAASPRLES 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2318	100.0	451	10	US-09-860-797-2	Sequence 2, Appli
2	2318	100.0	451	10	US-09-740-033-2	Sequence 2, Appli
3	821.5	35.4	419	10	US-09-740-033-4	Sequence 4, Appli
4	438.5	18.9	368	10	US-09-768-703-2	Sequence 2, Appli
5	349	15.1	359	10	US-09-989-861-19	Sequence 19, Appl
6	341.5	14.7	439	10	US-09-829-631A-13	Sequence 13, Appl
7	340	14.7	359	10	US-09-951-622-13	Sequence 13, Appl
8	329.5	14.2	437	10	US-09-829-631A-8	Sequence 8, Appli
9	326.5	14.1	400	10	US-09-895-211-4	Sequence 4, Appli
10	326.5	14.1	400	10	US-09-895-211-6	Sequence 6, Appli
11	320	13.8	501	10	US-09-951-622-9	Sequence 9, Appli
12	319.5	13.8	394	10	US-09-993-844-7	Sequence 7, Appli
13	319.5	13.8	408	10	US-09-895-211-5	Sequence 5, Appli
14	309.5	13.4	408	10	US-09-895-211-2	Sequence 2, Appli
15	308	13.3	517	10	US-09-951-622-10	Sequence 10, Appl
16	305	13.2	497	12	US-10-052-589-2	Sequence 2, Appli
17	300	12.9	466	10	US-09-951-622-11	Sequence 11, Appl
18	291.5	12.6	388	10	US-09-989-861-8	Sequence 8, Appli
19	286.5	12.4	529	10	US-09-951-622-2	Sequence 2, Appli
20	286	12.3	259	10	US-09-796-338A-23	Sequence 23, Appl
21	279.5	12.1	451	10	US-09-993-844-6	Sequence 6, Appli
22	271.5	11.7	291	10	US-09-829-631A-10	Sequence 10, Appl
23	270	11.6	445	10	US-09-989-861-16	Sequence 16, Appl
24	267.5	11.5	382	10	US-09-993-844-5	Sequence 5, Appli
25	264.5	11.4	222	10	US-09-911-005-4	Sequence 4, Appli
26	263	11.3	387	10	US-09-989-861-2	Sequence 2, Appli
27	263	11.3	406	10	US-09-989-861-4	Sequence 4, Appli
28	259	11.2	460	10	US-09-989-861-18	Sequence 18, Appl
29	250.5	10.8	390	10	US-09-160-116-19	Sequence 19, Appl
30	250	10.8	431	10	US-09-826-508-22	Sequence 22, Appl
31	250	10.8	431	10	US-09-899-532-2	Sequence 2, Appli
32	247	10.7	601	10	US-09-782-980-84	Sequence 84, Appl
33	247	10.7	601	10	US-09-884-430-8	Sequence 8, Appli
34	243	10.5	505	9	US-10-029-009-21	Sequence 21, Appl
35	241.5	10.4	428	10	US-09-292-973-4	Sequence 4, Appli
36	241	10.4	488	9	US-10-029-009-9	Sequence 9, Appli
37	239	10.3	518	10	US-09-804-551B-8	Sequence 8, Appli
38	238.5	10.3	481	8	US-08-681-219-31	Sequence 31, Appl
39	238.5	10.3	481	10	US-09-919-497-72	Sequence 72, Appl
40	236	10.2	398	10	US-09-796-338A-17	Sequence 17, Appl
41	236	10.2	398	10	US-09-897-201-2	Sequence 2, Appli
42	234	10.1	402	12	US-10-077-874-2	Sequence 2, Appli
43	234	10.1	425	10	US-09-961-848-2	Sequence 2, Appli
44	233.5	10.1	427	10	US-09-730-931-2	Sequence 2, Appli
45	233	10.1	370	12	US-10-044-592-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-860-797-2
; Sequence 2, Application US/09860797
; Patent No. US20020064825A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc.
; APPLICANT: Marilyn Evelyn Lewis
; APPLICANT: Nicola Melanie Robas
; TITLE OF INVENTION: No. US20020064825A1el Polypeptide
; FILE REFERENCE: PCS10924APME
; CURRENT APPLICATION NUMBER: US/09/860,797
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: GB 0012248.1
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: GB 0107394.9
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/211,421
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/283,441
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-860-797-2

Query Match 100.0%; Score 2318; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 3.6e-189;
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
          |||||||||||||||||||||||||||||||
Db    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451


RESULT 2
US-09-740-033-2
; Sequence 2, Application US/09740033
; Patent No. US20020100067A1
; GENERAL INFORMATION:
;  APPLICANT: GAN, Weiniu et al
;  TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
;  TITLE OF INVENTION:  RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
;  TITLE OF INVENTION:  PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL001056
;  CURRENT APPLICATION NUMBER: US/09/740,033
;  CURRENT FILING DATE:  2000-12-20
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 451
;   TYPE: PRT
;   ORGANISM: Human
US-09-740-033-2

  Query Match            100.0%;  Score 2318;  DB 10;  Length 451;
  Best Local Similarity  100.0%;  Pred. No. 3.6e-189;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy     61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy    121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
          |||||||||||||||||||||||||||||||
Db    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451


RESULT 3
US-09-740-033-4
; Sequence 4, Application US/09740033
; Patent No. US20020100067A1
; GENERAL INFORMATION:
;  APPLICANT: GAN, Weiniu et al
;  TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
;  TITLE OF INVENTION:  RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
;  TITLE OF INVENTION:  PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL001056
;  CURRENT APPLICATION NUMBER: US/09/740,033
;  CURRENT FILING DATE:  2000-12-20
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Oryzias latipes
US-09-740-033-4

  Query Match            35.4%;  Score 821.5;  DB 10;  Length 419;
  Best Local Similarity  42.8%;  Pred. No. 3e-62;
  Matches  182;  Conservative   78;  Mismatches  140;  Indels   25;  Gaps   10;

Qy      9 SSGNSST--LGRVPQTPGPSTASGVPEVGL----RDVASESVALFFMLLLDLTAVAGNAA 62
          |  | ||   |  |  |        |:||:    :    :   || |: |:| |:  |
Db      8 SISNFSTGLFGPHPTVP--------PDVGVVTSSQSQMKDLFGLFCMVTLNLIALLANTG 59

Qy     63 VMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLS 122
          ||  ||: | |:|| || ||| ||:| |: |||| ::||| |   :|  : |::|:||:
Db     60 VMVAIARAPHLKKFAFVCHLCAVDVLCAILLMPLGIISSSPFFGTVVFTILECQVYIFLN 119

Qy    123 VCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLGRV 182
          |  : |:||:::||:||||:|:|||||||:||: ||  |:: :| |:| :| | : |
Db    120 VFLIWLSILTITAISVERYFYIVHPMRYEVKMTINLVIGVMLLIWFKSLLLALVTLFGWP 179

Qy    183 SWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQH 242
          :   :      |||  |||    :| |:| |: || |:::|  || ::::|||| ||:|
Db    180 PYGHQSSIAASHCSLHASHSRLRGVFAVLFCVICFLAPVVVIFSVYSAVYKVARSAALQQ 239

Qy    243 GP-LPTWMET--PRQRSESLSSRSTMVTSSGAPQ-TTPHRTFGGGKAAVVLLAVGGQFLL 298
          | :||| :    : ||:|::|::|::|:   ||  :| | | |||||: |  : ||||:
Db    240 VPAVPTWADASPAKDRSDSINSQTTIITTRTLPQRLSPERAFSGGKAALTLAFIVGQFLV 299

Qy    299 CWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSK-QFV 357
          ||||:| ||| ::|:    | | :|  | |: |  |  || ||| ||||||  || | :
Db    300 CWLPFFIFHLQMSLTGSMKSPGDLEEAVNWLAYSSFAVNPSFYGLLNRQIRDELVKFRRC 359

Qy    358 CFFKPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPK-QEPPAVDFRIPG 416
          |  :|    |:  | ||| :||||||:| |   ||:  |    :|:  |  |   :|||
Db    360 CVTQPV---EIGPSSLEGSFQENFLQFIQRTSSSSETHPSFANSNPRNMENQA--HKIPG 414

Qy    417 QIAEE 421
          || ||
Db    415 QIPEE 419


RESULT 4
US-09-768-703-2
; Sequence 2, Application US/09768703
; Patent No. US20020098538A1
; GENERAL INFORMATION:
;  APPLICANT: SHABON, USMAN
;  APPLICANT:  ELSHOURBAGY, NABIL
;  APPLICANT:  MICHALOVICH, DAVID
;  TITLE OF INVENTION: 7TM RECEPTOR (AXOR23)
```

```
;  FILE REFERENCE: GP-30197 C1
;  CURRENT APPLICATION NUMBER: US/09/768,703
;  CURRENT FILING DATE:  2001-01-24
;  PRIOR APPLICATION NUMBER: UK 9905317.5
;  PRIOR FILING DATE: 1999-03-09
;  PRIOR APPLICATION NUMBER: 09/396,610
;  PRIOR FILING DATE: 1999-09-15
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 368
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-09-768-703-2

  Query Match             18.9%;  Score 438.5;  DB 10;  Length 368;
  Best Local Similarity   31.8%;  Pred. No. 7.9e-30;
  Matches  127;  Conservative   55;  Mismatches  146;  Indels   71;  Gaps   14;

Qy      40 ASE---SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPL 96
           |||   |: |    :::: |: || |:: |: :|| ||  ::: |||:|||||| ::|||
Db       9 ASEVAGSLGLILAAVVEVGALLGNGALLVVVLRTPGLRDALYLAHLCVVDLLAAASIMPL 68

Qy      97 AMLSSSAL-FDHALFGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMT 155
            :|::          |   ||   |||   :    | |:|: : ||  :|||:|   |
Db      69 GLLAAPPPGLGRVRLGPAPCRAARFLSAALLPACTLGVAALGLARYRLIVHPLRPGSRPP 128

Qy     156 LGLVASVLVGVWVKALAMASVPVLGRVSWEEGAPSVPPGCSLQWSHSAYCQL-------F 208
             |   ||  ||  |  : :: :|       | |  ||         | | :       |
Db     129 PVL---VLTAVWAAAGLLGALSLL-------GPPPAPPPA------PARCSVLAGGLGPF 172

Qy     209 VVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTS 268
             ::|:| | || ||:|  |  :| ||| ||::  | |      | ||:|| || :::
Db     173 RPLWALLAFALPALLLLGAYGGIFVVARRAALR-PPRPA--RGSRLRSDSLDSRLSIL-- 227

Qy     269 SGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTW 328
             |   |     |||||:      |||  |||||       |  |      : |: |||
Db     228 ---PPLRPR--LPGGKAALAPALAVGQFAACWLPY-----GCACLAPAARAAEAEAAVTW 277

Qy     329 IGYFCFTSNPFFYGCLNRQIR---GELSKQFVCFFKPAPEEELRLPSREGSIEENFLQFL 385
           : |  | ::|| || | | :|   | ||:: :    | |
Db     278 VAYSAFAAHPFLYGLLQRPVRLALGRLSRRAL----PGPVR------------------- 314

Qy     386 QGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAEETSE 424
             |  ::|  | |    | ||      | :  |:| |
Db     315 ---ACTPQAWHPRALLQCLQRPPEGPAVGPSEAPEQTPE 350


RESULT 5
US-09-989-861-19
; Sequence 19, Application US/09989861
; Patent No. US20020081661A1
; GENERAL INFORMATION:
;  APPLICANT: Gerald, Christophe
;  APPLICANT:  Hartig, Paul R.
;  APPLICANT:  Branchek, Theresa
;  APPLICANT:  Weinshank, Richard L.
;  TITLE OF INVENTION: DNA Encoding 5-HT4 Serotonin Receptors And Uses
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: 42667-A2-PCT-US
;  CURRENT APPLICATION NUMBER: US/09/989,861
;  CURRENT FILING DATE:  2001-11-19
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/328,314
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-03
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US93/12586
;  PRIOR FILING DATE: EARLIER FILING DATE: 1993-12-22
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 07/996,772
;  PRIOR FILING DATE: EARLIER FILING DATE: 1992-12-24
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: canine
US-09-989-861-19

  Query Match             15.1%;  Score 349;  DB 10;  Length 359;
  Best Local Similarity   28.7%;  Pred. No. 3e-22;
  Matches  100;  Conservative   56;  Mismatches  140;  Indels   52;  Gaps   10;

Qy      43 SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLS 100
           || |  ::|:    :|||  |   :     ||     |:  | : |||  | ::|
Db      22 SVVLTVLILI---TIAGNVVVCLAVGLNRRLRSLTNCFIVSLAITDLLLGLLVLPF---- 74

Qy     101 SSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLG 157
            ||  :   :    ||:| | :|  | |   : :||::  |:::||  |  |:|| | :|
Db      75 -SAFYQLSCRWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPV 133

Qy     158 LVASVLVGVWVKALAMASVPV-LGRVSWEEGAP--SVPPGCSLQWSHSAYCQLFVVVFAV 214
            ||  || :|| :: :: : : ||  |  | :       | | :| :       :: :|  :
Db     134 RVAVSLVLIWVISITLSFLSIHLGWNSRNETSSFNHTIPKCKVQVN-----LVYGLVDGL 188

Qy     215 LYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQT 274
           : | ||||:: : |  :|::||  | :   : :|
Db     189 VTFYLPLLVMCITYYRIFKIARDQAKRIHHMGSWKAA---------------------- 225

Qy     275 TPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCF 334
             | |  || | | || | |::|| |||:  :|  |           |:|| |:||
Db     226 ----TIGEHKATVTLAAVMGAFIICWFPYFTVFVYRGLKGDDAINEAFEAVVLWLGYANS 281

Qy     335 TSNPFFYGCLNRQIRGELSKQFVCFFKPAPE--EELRLPSREGSIEEN 380
             ||  |  |||  |    : | |  :||    :|  | |    :  |
Db     282 ALNPILYATLNRDFRTAYQQLFRC--RPASHNAQETSLRSNSSQLARN 327


RESULT 6
US-09-829-631A-13
; Sequence 13, Application US/09829631A
; Patent No. US20020091235A1
; GENERAL INFORMATION:
;  APPLICANT: Sibley, David R.
;  APPLICANT:  Monsma, Frederick J.
;  APPLICANT:  Hamblin, Mark
;  TITLE OF INVENTION: The ST-B17 Serotonin Receptor
;  FILE REFERENCE: NIH047.1CP1C1
;  CURRENT APPLICATION NUMBER: US/09/829,631A
;  CURRENT FILING DATE:  2001-04-10
;  PRIOR APPLICATION NUMBER: US 08/428,242
;  PRIOR FILING DATE: 1995-09-18
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 439
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(439)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-829-631A-13

  Query Match             14.7%;  Score 341.5;  DB 10;  Length 439;
  Best Local Similarity   29.8%;  Pred. No. 1.7e-21;
  Matches  127;  Conservative   64;  Mismatches  182;  Indels   53;  Gaps   17;

Qy      19 VPQTPGPSTASGVPE--VGLRDVASES-VALFFMLLLDLTAVAGNAAVMAVIAKTPALRK 75
           ||: ||| ||:  |    | |       ||    ::: ||| | |: ::|:|   ||||
Db       2 VPE-PGP-TANSTPAWGAGARXXGGSGWVAAGLCVVIALTA-AANSLLIALICTQPALRN 58

Qy      76 F--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILSV 133
           |:  |   ||:  | :|| ||| :||:   :     | |:    |   | :||::
Db      59 TSNFFLVSLFTSDLMVGLVVMPPAML--NALYGRWVLARGLCLLWTAFDVMCCSASILNL 116

Qy     134 SAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLGRVSWEE---GAPS 190
            |:::||  :: |:||::|||     ::::| |  |   : :|:|  | |     |
Db     117 CLISLDRYLLILSPLRYKLRMTPLRALALVLGGWSLAALASFLPLL--LGWHELGHARPP 174
```

```
Qy    191 VPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWME 250
          ||  | |  |       ||:| : | | ||   |   || :   ||  |:|   | | |
Db    175 VPGQCRLLASLP-----FVLVASGLTFFLPSGAICFTYCRILLAARKQAVQVASLTTGMA 229

Qy    251 TPRQRSESLS---SRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFH 307
          :  | ||:|    | :  | |: : :     :  | ||:: |  : | | : |||:|  :
Db    230 S--QASETLQVPRSPAAGVESADSRRLATKSSRKGLKASMTLGILLGMFFVTWLPFFVAN 287

Qy    308 LYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFFKP----- 362
          :  |:    || |  : |:||:||   | ||  |       :  | :   |   |
Db    288 IVQAV-CDCISPGLFD-VLTWLGYCNSTMNPIIYPLFMLDFKRALGRFLPCPRCPREPRP 345

Qy    363 -----------APEEEL------RLPSREGSIEENFLQFLQGTGC---PSESWVSRPLP 401
                      ||   |      | | |   |             |   ||  :::||
Db    346 AWPRHHCAPLTAAPGPALAYSRCCRCPCRRTQIRTQTQAQAAPRACGSRPSCCFLARPPR 405

Qy    402 SPKQEP 407
          :|:  |
Db    406 TPRCPP 411


RESULT 7
US-09-951-622-13
; Sequence 13, Application US/09951622
; Patent No. US20020106734A1
; GENERAL INFORMATION:
;  APPLICANT: Daniel R. Soppet et al.
;  TITLE OF INVENTION: ADRENERGIC RECEPTOR
;  FILE REFERENCE: PF128D2C1
;  CURRENT APPLICATION NUMBER: US/09/951,622
;  CURRENT FILING DATE:  2001-09-14
;  PRIOR APPLICATION NUMBER: 09/339,244
;  PRIOR FILING DATE: 1999-06-24
;  PRIOR APPLICATION NUMBER: 09/030,582
;  PRIOR FILING DATE: 1998-02-25
;  PRIOR APPLICATION NUMBER: 08/467,568
;  PRIOR FILING DATE: 1995-06-06
;  PRIOR APPLICATION NUMBER: PCT/US94/09051
;  PRIOR FILING DATE: 1994-08-10
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
;  LENGTH: 359
;  TYPE: PRT
;  ORGANISM: human
US-09-951-622-13

  Query Match             14.7%;  Score 340;  DB 10;  Length 359;
  Best Local Similarity   28.0%;  Pred. No. 1.7e-21;
  Matches   97;  Conservative   57;  Mismatches  134;  Indels   58;  Gaps    9;

Qy     49 MLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFD 106
          : :| |  ||||  |   :     ||     |:  | : |||  | ::|      ||::
Db     25 LAVLILITVAGNVVVCLAVGLNRRLRNLTNCFIVSLAITDLLLGLLVLPF-----SAIYQ 79

Qy    107 HAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVL 163
          :    ||:| | :|  | |   : :|||::  |:::||  |: |:|| | :|   ||  |
Db     80 LSCKWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVMDPLRYPVLVTPVRVAISL 139

Qy    164 VGVWVKALAMASVPV-LGRVSWEEGAPS--VPPGCSLQWSHSAYCQLFVVVFAVLYFLLP 220
          | :|| :: :: : : ||  |  | :       | :| :     ::: :|  :: | ||
Db    140 VLIWVISITLSFLSIHLGWNSRNETSKGNHTTSKCKVQVN-----EVYGLVDGLVTFYLP 194

Qy    221 LLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTF 280
          ||:: : |  :|:|||  | :   : :|                                |
Db    195 LLIMCITYYRIFKVARDQAKRINHISSWKAA--------------------------TI 227

Qy    281 GGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFF 340
             || | | || | |::|| |||:  :|  |          :|::| |:||     ||
Db    228 REHKATVTLAAVMGAFIICWFPYFTAFVYRGLRGDDAINEVLEAIVLWLGYANSALNPIL 287

Qy    341 YGCLNRQIRGELSKQFVCFFKPAPEEELRLPSREG---SIEENFLQ 383


          |  |||  |    : | |          || :|     |:  |  |
Db    288 YAALNRDFRTGYQQLFCC----------RLANRNSHKTSLRSNASQ 323


RESULT 8
US-09-829-631A-8
; Sequence 8, Application US/09829631A
; Patent No. US20020091235A1
; GENERAL INFORMATION:
;  APPLICANT: Sibley, David R.
;  APPLICANT:  Monsma, Frederick J.
;  APPLICANT:  Hamblin, Mark
;  TITLE OF INVENTION: The ST-B17 Serotonin Receptor
;  FILE REFERENCE: NIH047.1CP1C1
;  CURRENT APPLICATION NUMBER: US/09/829,631A
;  CURRENT FILING DATE:  2001-04-10
;  PRIOR APPLICATION NUMBER: US 08/428,242
;  PRIOR FILING DATE: 1995-09-18
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;  LENGTH: 437
;  TYPE: PRT
;  ORGANISM: Rat
US-09-829-631A-8

  Query Match             14.2%;  Score 329.5;  DB 10;  Length 437;
  Best Local Similarity   27.3%;  Pred. No. 1.7e-20;
  Matches  128;  Conservative   65;  Mismatches  181;  Indels   95;  Gaps   19;

Qy      2 ESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNA 61
            |  |:     ||||       |||  | |          |  ||    ::: ||| | |:
Db      4 EPGPV-----NSST---PAWGPGPPPAPG---------GSGWVAAALCVVIVLTA-AANS 45

Qy     62 AVMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYL 119
             :: :|   ||:|     |:  |   ||:  | :|| |||  :||:   :     | |:
Db     46 LLIVLICTQPAVRNTSNFFLVSLFTSDLMVGLVVMPPAML--NALYGRWVLARGLCLLWT 103

Qy    120 FLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVL 179
             |   | :||::  |:::||  :: |:|||::|||     ::::| |  |   : :|:|
Db    104 AFDVMCCSASILNLCLISLDRYLLILSPLRYKLRMTAPRALALILGAWSLAALASFLPLL 163

Qy    180 GRVSWEE-GAPSVP-PGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMF 232
            : | | |     | ||          |:|     ||:| : : | ||   |   || :
Db    164 --LGWHELGKARTPAPG---------QCRLLASLPFVLVASGVTFFLPSGAICFTYCRIL 212

Qy    233 RVARVAAMQHGPLPTWM----------ETPRQRSESLSSRSTMVTSSGAPQTTPHRTFG 281
          ||  |:|   | |              |||   ||  ||       |
Db    213 LAARKQAVQVASLTTGTAGQALETLQVPRTPRPGMESADSRRLATKHSRKAL-------- 264

Qy    282 GGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFY 341
          ||:: |  : | | : |||:|  ::  |:     || |  : |:||:||   | ||  |
Db    265 --KASLTLGILLGMFFVTWLPFFVANIAQAV-CDCISPGLFD-VLTWLGYCNSTMNPIIY 320

Qy    342 GCLNRQIRGELSKQFV----------------------CFFKPAPEEELRLP------S 372
             |  :  | : |:                        |  :|  :: | ||       |
Db    321 PLFMRDFKRALGR-FLHASTVPRSTGQPCLPLHVDLSQRCQTRPQLQQVLALPLPPNSDS 379

Qy    373 REGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAEE 421
             |   : ||       | |:  :|  | | :    |:| :   :  |
Db    380 DSASGGTSGLQLTAQLLLPGEA--TRDPPPPTRATTVVNFFVTDSVEPE 426


RESULT 9
US-09-895-211-4
; Sequence 4, Application US/09895211
; Patent No. US20020127639A1
; GENERAL INFORMATION:
;  APPLICANT: Hunton and Williams
;  APPLICANT:  Emorine, Laurent
;  TITLE OF INVENTION: INTRON/EXON STRUCTURE OF THE HUMAN AND MOUSE BETA3 ADRENERGIC
;  TITLE OF INVENTION:  GENES
;  FILE REFERENCE: 58769.000011
```

```
;  CURRENT APPLICATION NUMBER: US/09/895,211
;  CURRENT FILING DATE:  2001-07-02
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-895-211-4

  Query Match             14.1%;  Score 326.5;  DB 10;  Length 400;
  Best Local Similarity   27.4%;  Pred. No. 2.8e-20;
  Matches  107;  Conservative   61;  Mismatches  171;  Indels   51;  Gaps   12;

Qy       4 SPIPQSSGNSSTLGRVPQ-TPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAA 62
Db       2 APWPHRNGSLALWSDAPTLDPSAANTSGLPGV-------PWAAALAGALLALATVGGNLL 54

Qy      63 VMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
Db      55 VIIAIARTPRLQTITNVFVTSLAAADLVVGLLVMPPG--ATLALTGHWPLGETGCELWTS 112

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
Db     113 VDVLCVTASIETLCALAVDRYLAVTNPLRYGTLVTKRRARAAVVLVWIVSAAVSFAPIMS 172

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA 235
Db     173 Q-WWRVGADAEAQECH---SNPRCCSFASNMPYALLSSSVSFYLPLLVMLFVYARVFVVA 228

Qy     236 ------------RVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAP-------QTTP 276
Db     229 KRQRHLLRRELGRFSPEESPPSPS-------RSPSPATGGTPAAPDGVPPCGRRPARLLP 281

Qy     277 HRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTS 336
Db     282 LREH---RALRTLGLIMGIFSLCWLPFFLANVLRALAGPSLVPSGVFIALNWLGYANSAF 338

Qy     337 NPFFYGCLNRQIRGELSKQFVCFFKPAPEE 366
Db     339 NPVIY-CRSPDFRDAFRRLLCSYGGRGPEE 367


RESULT 10
US-09-895-211-6
; Sequence 6, Application US/09895211
; Patent No. US20020127639A1
; GENERAL INFORMATION:
;  APPLICANT: Hunton and Williams
;  APPLICANT:  Emorine, Laurent
;  TITLE OF INVENTION: INTRON/EXON STRUCTURE OF THE HUMAN AND MOUSE BETA3 ADRENERGIC REC
;  TITLE OF INVENTION:  GENES
;  FILE REFERENCE: 58769.000011
;  CURRENT APPLICATION NUMBER: US/09/895,211
;  CURRENT FILING DATE:  2001-07-02
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-895-211-6

  Query Match             14.1%;  Score 326.5;  DB 10;  Length 400;
  Best Local Similarity   27.4%;  Pred. No. 2.8e-20;
  Matches  107;  Conservative   61;  Mismatches  171;  Indels   51;  Gaps   12;

Qy       4 SPIPQSSGNSSTLGRVPQ-TPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAA 62
Db       2 APWPHRNGSLALWSDAPTLDPSAANTSGLPGV-------PWAAALAGALLALATVGGNLL 54

Qy      63 VMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
Db      55 VIIAIARTPRLQTITNVFVTSLAAADLVVGLLVMPPG--ATLALTGHWPLGETGCELWTS 112

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
Db     113 VDVLCVTASIETLCALAVDRYLAVTNPLRYGTLVTKRRARAAVVLVWIVSAAVSFAPIMS 172

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA 235
Db     173 Q-WWRVGADAEAQECH---SNPRCCSFASNMPYALLSSSVSFYLPLLVMLFVYARVFVVA 228

Qy     236 ------------RVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAP-------QTTP 276
Db     229 KRQRHLLRRELGRFSPEESPPSPS-------RSPSPATGGTPAAPDGVPPCGRRPARLLP 281

Qy     277 HRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTS 336
Db     282 LREH---RALRTLGLIMGIFSLCWLPFFLANVLRALAGPSLVPSGVFIALNWLGYANSAF 338

Qy     337 NPFFYGCLNRQIRGELSKQFVCFFKPAPEE 366
Db     339 NPVIY-CRSPDFRDAFRRLLCSYGGRGPEE 367


RESULT 11
US-09-951-622-9
; Sequence 9, Application US/09951622
; Patent No. US20020106734A1
; GENERAL INFORMATION:
;  APPLICANT: Daniel R. Soppet et al.
;  TITLE OF INVENTION: ADRENERGIC RECEPTOR
;  FILE REFERENCE: PF128D2C1
;  CURRENT APPLICATION NUMBER: US/09/951,622
;  CURRENT FILING DATE:  2001-09-14
;  PRIOR APPLICATION NUMBER: 09/339,244
;  PRIOR FILING DATE: 1999-06-24
;  PRIOR APPLICATION NUMBER: 09/030,582
;  PRIOR FILING DATE: 1998-02-25
;  PRIOR APPLICATION NUMBER: 08/467,568
;  PRIOR FILING DATE: 1995-06-06
;  PRIOR APPLICATION NUMBER: PCT/US94/09051
;  PRIOR FILING DATE: 1994-08-10
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;   LENGTH: 501
;   TYPE: PRT
;   ORGANISM: human
US-09-951-622-9

  Query Match             13.8%;  Score 320;  DB 10;  Length 501;
  Best Local Similarity   26.3%;  Pred. No. 1.3e-19;
  Matches  113;  Conservative   67;  Mismatches  167;  Indels   82;  Gaps   15;

Qy      26 STASGVPEVGLRDVASESVAL-FFMLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHL 82
Db      41 STAA----VGGLVVSAQGVGVGVFLAAFILMAVAGNLLVILSVACNRHLQTVTNYFIVNL 96

Qy      83 CLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILSVSAINVERYY 142
Db      97 AVADLLLSATVLPFS--ATMEVLGFWAFGRAFCDVWAAVDVLCCTASILSLCTISVDRYV 154

Qy     143 YVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLGRVSWEEGAPSVPPGCSLQWSHS 202
Db     155 GVRHSLKYPAIMTERKAAAILALLWVVALVVSVGPLLG---WKE---PVPP-------DE 201

Qy     203 AYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVAR--VAAMQHGPLPTWMETPRQR 255
Db     202 RFCGITEEAGYAVFSSVCSFYLPMAVIVVMYCRVYVVARSTTRSLEAG-------VKRER 254

Qy     256 SES------LSSRSTMVTSSGA--PQTTPHRTFGGG------------KAAVVLLAVGGQ 295
Db     255 GKASEVVLRIHCRGAATGADGAHGMRSAKGHTFRSSLSVRLLKFSREKKAAKTLAIVVGV 314
```

```
Qy    296 FLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQ 355
          |:||| |:|      :|  |   :  | |: |:|||    ||  | | :|: :    :
Db    315 FVLCWFPFFFVLPLGSLFPQLKPSEGVFKVIFWLGYFNSCVNPLIYPCSSREFKRAFLRL 374

Qy    356 FVCFFK----------------PAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRP 399
            |  :                |    ||      |          |   |
Db    375 LRCQCRRRRRRRPLWRVYGHHWRASTSGLRQDCAPSS----------GDAPPGAPLALTA 424

Qy    400 LPSPKQEPP 408
          || |  |||
Db    425 LPDPDPEPP 433
```

```
RESULT 12
US-09-993-844-7
; Sequence 7, Application US/09993844
; Patent No. US20020106739A1
; GENERAL INFORMATION:
;  APPLICANT: Oakley, Robert H.
;  APPLICANT:  Barak, Lawrence S.
;  APPLICANT:  Laporte, Stephane A.
;  APPLICANT:  Caron, Marc G.
;  TITLE OF INVENTION: Modified G-Protein Coupled Receptors
;  FILE REFERENCE: 033072-026
;  CURRENT APPLICATION NUMBER: US/09/993,844
;  CURRENT FILING DATE:  2001-11-05
;  PRIOR APPLICATION NUMBER: US 60/245,772
;  PRIOR FILING DATE: 2000-11-03
;  PRIOR APPLICATION NUMBER: US 60/260,363
;  PRIOR FILING DATE: 2001-01-08
;  NUMBER OF SEQ ID NOS: 81
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 394
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence of beta3AR-V2R chimera
;   OTHER INFORMATION: expressed from pEArrB-1/beta3AR vector
US-09-993-844-7

  Query Match             13.8%;  Score 319.5;  DB 10;  Length 394;
  Best Local Similarity   29.3%;  Pred. No. 1.1e-19;
  Matches  111;  Conservative   61;  Mismatches  164;  Indels   43;  Gaps   16;

Qy      4 SPIPQSSGNSSTLGRVPQTP--GPSTA--SGVPEVGLRDVASESVALFFMLLLDLTAVAG 59
          :| |     :|:|   |  |   |:|| ||:| |        ::|   : |  |  | |
Db      2 APWPH---ENSSLAPWPDLPTLAPNTANTSGLPGVPW----EAALAGALLALAVLATVGG 54

Qy     60 NAAVMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRL 117
          |  |:  || || |:    ||| |   ||:  | ::| |  :: ||  |   |  | |
Db     55 NLLVIVAIAWTPRLQTMTNVFVTSLAAADLVMGLLVVPPA--ATLALTGHWPLGATGCEL 112

Qy    118 YLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVP 177
          :  : |  |: :| :: |: |:||  | :|:||   :|     : :| ||| : |::  |
Db    113 WTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALVTKRCARTAVVLVWVVSAAVSFAP 172

Qy    178 VLGRVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMF 232
          :: :  |  || :    |    |:   |       :|:: : : | ||||:|  ||  :|
Db    173 IMSQ-WWRVGADAEAQRCH---SNPRCCAFASNMPYVLLSSSVSFYLPLLVMLFVYARVF 228

Qy    233 RVA----RVAAMQHGPLPTWMETPRQRSESLSSR--STMVTSSGAP-------QTTPHRT 279
          ||    |:   : |  |   |:|   | ||:     |      | |        :  | |
Db    229 VVATRQLRLLRGELGRFPP-EESPPAPSRSLAPAPVGTCAPPEGVPACGRRPARLLPLRE 287

Qy    280 FGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPF 339
            :|   | : | | ||||:|  ::  ||    :  |     : |:||     ||
Db    288 H---RALCTLGLIMGTFTLCWLPFFLANVLRALGGPSLVPGPAFLALNWLGYANSAFNPL 344

Qy    340 FYGCLNRQIRGELSKQFVC 358
          | | :   |    :: :|
Db    345 IY-CRSPDFRSAF-RRLLC 361
```

```
RESULT 13
US-09-895-211-5
; Sequence 5, Application US/09895211
; Patent No. US20020127639A1
; GENERAL INFORMATION:
;  APPLICANT: Hunton and Williams
;  APPLICANT:  Emorine, Laurent
;  TITLE OF INVENTION: INTRON/EXON STRUCTURE OF THE HUMAN AND MOUSE BETA3 ADRENERGIC
;  TITLE OF INVENTION:  GENES
;  FILE REFERENCE: 58769.000011
;  CURRENT APPLICATION NUMBER: US/09/895,211
;  CURRENT FILING DATE:  2001-07-02
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 408
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-895-211-5

  Query Match             13.8%;  Score 319.5;  DB 10;  Length 408;
  Best Local Similarity   29.3%;  Pred. No. 1.1e-19;
  Matches  111;  Conservative   61;  Mismatches  164;  Indels   43;  Gaps   16;

Qy      4 SPIPQSSGNSSTLGRVPQTP--GPSTA--SGVPEVGLRDVASESVALFFMLLLDLTAVAG 59
          :| |     :|:|   |  |   |:|| ||:| |        ::|   : |  |  | |
Db      2 APWPH---ENSSLAPWPDLPTLAPNTANTSGLPGVPW----EAALAGALLALAVLATVGG 54

Qy     60 NAAVMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRL 117
          |  |:  || || |:    ||| |   ||:  | ::| |  :: ||  |   |  | |
Db     55 NLLVIVAIAWTPRLQTMTNVFVTSLAAADLVMGLLVVPPA--ATLALTGHWPLGATGCEL 112

Qy    118 YLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVP 177
          :  : |  |: :| :: |: |:||  | :|:||   :|     : :| ||| : |::  |
Db    113 WTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALVTKRCARTAVVLVWVVSAAVSFAP 172

Qy    178 VLGRVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMF 232
          :: :  |  || :    |    |:   |       :|:: : : | ||||:|  ||  :|
Db    173 IMSQ-WWRVGADAEAQRCH---SNPRCCAFASNMPYVLLSSSVSFYLPLLVMLFVYARVF 228

Qy    233 RVA----RVAAMQHGPLPTWMETPRQRSESLSSR--STMVTSSGAP-------QTTPHRT 279
          ||    |:   : |  |   |:|   | ||:     |      | |        :  | |
Db    229 VVATRQLRLLRGELGRFPP-EESPPAPSRSLAPAPVGTCAPPEGVPACGRRPARLLPLRE 287

Qy    280 FGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPF 339
            :|   | : | | ||||:|  ::  ||    :  |     : |:||     ||
Db    288 H---RALCTLGLIMGTFTLCWLPFFLANVLRALGGPSLVPGPAFLALNWLGYANSAFNPL 344

Qy    340 FYGCLNRQIRGELSKQFVC 358
          | | :   |    :: :|
Db    345 IY-CRSPDFRSAF-RRLLC 361
```

```
RESULT 14
US-09-895-211-2
; Sequence 2, Application US/09895211
; Patent No. US20020127639A1
; GENERAL INFORMATION:
;  APPLICANT: Hunton and Williams
;  APPLICANT:  Emorine, Laurent
;  TITLE OF INVENTION: INTRON/EXON STRUCTURE OF THE HUMAN AND MOUSE BETA3 ADRENERGIC
;  TITLE OF INVENTION:  GENES
;  FILE REFERENCE: 58769.000011
;  CURRENT APPLICATION NUMBER: US/09/895,211
;  CURRENT FILING DATE:  2001-07-02
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 408
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-895-211-2
```

```
  Query Match           13.4%;  Score 309.5;  DB 10;  Length 408;
  Best Local Similarity  29.4%;  Pred. No. 8e-19;
  Matches  112;  Conservative   60;  Mismatches  162;  Indels   47;  Gaps   17;

Qy       4 SPIPQSSGNSSTLGRVPQTP--GPSTA--SGVPEVGLRDVASESVALFFMLLLDLTAVAG 59
Db       2 APWPH---ENSSLAPWPDLPTLAPNTANTSGLP--GVPWEAALAGALLALAVLATTGV-- 54

Qy      60 NAAVMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRL 117
Db      55 NLLVIVAIAWTPRLQTMTNVFVTSLAAADLVMGLLVVPPA--ATLALTGHWPLGATGCEL 112

Qy     118 YLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVP 177
Db     113 WTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALVTKRCARTAVVLVWVVSAAVSFAP 172

Qy     178 VLGRVSWEEGAPSVP-------PGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCS 230
Db     173 IMSQ--WWRGVADAEAQRCHSNPRCCAFASQMPY----VLLSSSVSFYLPLLVMLFVYAR 226

Qy     231 MFRVA----RVAAMQHGPLPTWMETPRQRSESLSSR--STMVTSSGAP-------QTTPH 277
Db     227 VFVVATRQLRLLRGELGRFPP-EESPPAPSRSLAPAPVGTCAPPEGVPACGRRPARLLPL 285

Qy     278 RTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSN 337
Db     286 REH---RALCTLGLIMGTFTLCWLPFFLANVLRALGGPSLVPGPAFLALNWLGYANSAFN 342

Qy     338 PFFYGCLNRQIRGELSKQFVC 358
Db     343 PLIY-CRSPDFRSAF-RRLLC 361


RESULT 15
US-09-951-622-10
; Sequence 10, Application US/09951622
; Patent No. US20020106734A1
; GENERAL INFORMATION:
;  APPLICANT: Daniel R. Soppet et al.
;  TITLE OF INVENTION: ADRENERGIC RECEPTOR
;  FILE REFERENCE: PF128D2C1
;  CURRENT APPLICATION NUMBER: US/09/951,622
;  CURRENT FILING DATE:  2001-09-14
;  PRIOR APPLICATION NUMBER: 09/339,244
;  PRIOR FILING DATE: 1999-06-24
;  PRIOR APPLICATION NUMBER: 09/030,582
;  PRIOR FILING DATE: 1998-02-25
;  PRIOR APPLICATION NUMBER: 08/467,568
;  PRIOR FILING DATE: 1995-06-06
;  PRIOR APPLICATION NUMBER: PCT/US94/09051
;  PRIOR FILING DATE: 1994-08-10
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 517
;   TYPE: PRT
;   ORGANISM: human
US-09-951-622-10

  Query Match           13.3%;  Score 308;  DB 10;  Length 517;
  Best Local Similarity  25.9%;  Pred. No. 1.5e-18;
  Matches  117;  Conservative   84;  Mismatches  167;  Indels   84;  Gaps   21;

Qy      24 GP---STASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAAVMAVIAKTPALR--KFVF 78
Db      27 GPNQTSSNSTLPQLDITRAISVGLVLGAFILF---AIVGNILVILSVACNRHLRTPTNYF 83

Qy      79 VFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILSVSAINV 138
Db      84 IVNLAMADLLLSFTVLPFS--AALEVLGYWVLGRIFCDIWAAVDVLCCTASILSLCAISI 141

Qy     139 ERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLGRVSWEEGAPSVPPGCSLQ 198
Db     142 DRYIGVRYSLQYPTLVTRRKAILALLSVWVLSTVISIGPLLG---WKEPAPNDDKECGVT 198

Qy     199 WSHSAYCQLFVVVFAVL-YFLLPLLLILVVYCSMFRVAR--VAAMQHGPLPTWMETPRQ- 254
Db     199 E------EPFYALFSSLGSFYIPLAVILVMYCRVYIVAKRTTKNLEAGVMKE-MSNSKEL 251

Qy     255 --RSESLSSRSTMVTSSGAPQTTPHRTFG--------GGKAAVVLLAVGGQFLLCWLPYF 304
Db     252 TLRIHSKNFHEDTLSSTKAKGHNPRSSIAVKLFKFSREKKAAKTLGIVVGMFILCWLPFF 311

Qy     305 SFHLYVALSAQPIST--------GQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQF 356
Db     312 -----IAL---PLGSLFSTLKPPDAVFKVVFWLGYFNSCLNPIIYPCSSK----EFKRAF 359

Qy     357 V------CFFKPAPEEELRLP---------SREGSIE--ENFLQFLQGTG-CPSESWVSR 398
Db     360 VRILGCQCRGRRRRRRRRLGGCAYTYRPWTRGGSLERSQSRKDSLDDSGSCLSGSQRTL 419

Qy     399 PLPSP------KQEPPAV------DFRIPGQI 418
Db     420 PSASPSPGYLGRGAPPPVELCAFPEWKAPGAL 451


Search completed: December 10, 2002, 09:59:34
Job time : 13 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: December 10, 2002, 09:53:05 ; Search time 40 Seconds
(without alignments)
1502.401 Million cell updates/sec

Title: US-09-838-028-2
Perfect score: 2318
Sequence: 1 MESSPIPQSSGNSSTLGRVP.....SDIIMSDSYLRPAASPRLES 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2318	100.0	451	22	AAE12024	Human G-protein co
2	2318	100.0	451	22	AAU07900	Human serotonin-li
3	2318	100.0	451	22	AAU08710	Human histamine H2
4	2318	100.0	451	22	AAU04367	Human G-protein co
5	2318	100.0	451	22	AAG64123	Human G protein-co
6	2318	100.0	451	23	ABB81539	Human G-protein co
7	2318	100.0	451	23	AAU79732	Novel human GPCR,
8	2318	100.0	451	23	AAM50637	Human G protein co
9	2318	100.0	451	23	AAO14014	Human G protein-co
10	2318	100.0	451	23	AAU74909	Amino acid sequenc
11	2318	100.0	451	23	AAU11895	Human novel G prot
12	2318	100.0	451	23	AAE17073	Human G-protein co
13	2318	100.0	451	23	ABB05445	Human G-protein co
14	2318	100.0	869	22	AAU04386	GPCR-Gs fusion pro
15	2317	100.0	451	23	AAU11896	Human novel G prot
16	2313	99.8	451	22	AAU04382	Human G-protein co
17	2204	95.1	449	23	AAO14015	Rat G protein-coup
18	2185	94.3	449	23	AAM50638	Mouse G protein co
19	1472	63.5	279	22	AAG66453	Human G protein co
20	1276	55.0	252	17	AAW06595	G protein conjugat
21	1276	55.0	252	17	AAR91232	Rabbit G-protein c
22	821.5	35.4	419	23	ABB81540	Oryzias latipes G-
23	692	29.9	138	22	AAO05867	Human polypeptide
24	438.5	18.9	368	21	AAB19523	G protein coupled
25	438.5	18.9	368	21	AAY71293	Human orphan G pro
26	438.5	18.9	368	21	AAB02827	Human G protein co
27	438.5	18.9	368	22	AAU10303	G-protein coupled
28	438.5	18.9	368	22	AAG64288	Human GTP-binding
29	438.5	18.9	368	22	AAB70142	Human G protein-co
30	438.5	18.9	368	23	AAE19359	Human G-protein co
31	438.5	18.9	368	23	AAU75903	G-protein coupled
32	349.5	15.1	440	22	ABB56328	Non-endogenous hum
33	347	15.0	348	17	AAW00551	Human histamine H2
34	347	15.0	348	18	AAW31712	Protein sequence o
35	346	14.9	359	20	AAY00890	Canine histamine H
36	341	14.7	359	23	AAU10902	Human histamine H2
37	340	14.7	359	20	AAY00889	Human histamine H2
38	340	14.7	359	21	AAY57172	Sequence of HuH2R.
39	340	14.7	359	23	AAU10901	Human histamine H2
40	340	14.7	400	14	AAR32501	Beta-adrenergic re
41	335.5	14.5	405	15	AAR62515	Bovine beta3 adren
42	333	14.4	565	22	AAU05409	Human adrenorecept
43	332.5	14.3	572	15	AAR53071	Alpha 1a adrenergi
44	332.5	14.3	572	15	AAR58681	Human alpha-1A adr
45	332.5	14.3	572	16	AAR90039	Human alpha-1A adr

ALIGNMENTS

RESULT 1
AAE12024
ID AAE12024 standard; Protein; 451 AA.
XX
AC AAE12024;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human G-protein coupled receptor, GCREC-3.
XX
KW Human; G-protein coupled receptor; GCREC; gene therapy;
KW proteosome analysis; cell proliferative disorder; arteriosclerosis;
KW cancer; neurological disorder; Huntington's disease; Parkinson's disease;
KW cardiovascular disorder; atherosclerosis; congestive heart failure;
KW gastrointestinal disorder; gastritis; nausea; autoimmune; anaemia;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW metabolic disorder; diabetes; obesity; viral infection; drug screening;
KW chromosomal mapping; cytostatic; immunomodulatory; antiinflammatory;
KW virucide.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 117..137
FT /label= Transmembrane_domain
FT Domain 206..228
FT /label= Transmembrane_domain
FT Domain 288..306
FT /label= Transmembrane_domain
XX
PN WO200172836-A2.
XX
PD 04-OCT-2001.

XX
PF 29-MAR-2001; 2001WO-US10436.
XX
PR 29-MAR-2000; 2000US-193051P.
PR 06-APR-2000; 2000US-195155P.
PR 20-APR-2000; 2000US-199084P.
PR 28-APR-2000; 2000US-200551P.
PR 05-MAY-2000; 2000US-202278P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Policky JL, Tribouley CM, Tang YT, Baughn MR, Graul R, Khan FA;
PI Nguyen DB, Patterson C, Lal P, Au-young J, Yang J, Hafalia A;
PI Walia NK, Das D;
XX
DR WPI; 2001-616472/71.
DR N-PSDB; AAD19579.
XX
PT New human G-protein coupled receptors, useful for treatment and
PT diagnosis of e.g. cell proliferation, also screening for specific
PT modulators, and related nucleic acid -
XX
PS Claim 1; Page 104-105; 111pp; English.
XX
CC The present sequence is a human G-protein coupled receptor, GCREC-3
CC protein. The GCRECs are used for treating or preventing disorders
CC associated with decreased expression of functional GCREC, and for
CC identifying specific agonists and antagonists, also binding agents
CC and modulators. They can also be used for generating specific antibodies
CC and for proteosome analysis. Disorders that can be treated include
CC cell proliferative disorders, e.g., arteriosclerosis and cancer,
CC neurological disorders, e.g., Huntington's disease and Parkinson's
CC disease, cardiovascular disorders, e.g., atherosclerosis and congestive
CC heart failure, gastrointestinal disorders, e.g., gastritis and nausea,
CC autoimmune/inflammatory disorders, e.g., acquired immunodeficiency
CC syndrome (AIDS) and anaemia, metabolic disorders, e.g., diabetes and
CC obesity and viral infections. Nucleic acids that encode GCREC are
CC used for identifying agents that alter its expression, for assessing
CC toxicity of test compounds, and as sources of primers and probes for
CC diagnostic detection of GCREC DNA and of therapeutic antisense and
CC ribozyme sequences. They can also be used in gene therapy, for
CC chromosomal mapping, and for recombinant production of GCREC. The
CC antibodies are useful for diagnosis and monitoring of diseases
CC associated with GCREC expression, for detecting and purifying GCREC,
CC and as therapeutic agents and for drug screening.
XX
SQ Sequence 451 AA;

Query Match 100.0%; Score 2318; DB 22; Length 451;
Best Local Similarity 100.0%; Pred. No. 2.4e-228;
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy     61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy    121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
          |||||||||||||||||||||||||||||||
Db    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
```

RESULT 2
AAU07900
ID AAU07900 standard; Protein; 451 AA.
XX
AC AAU07900;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human serotonin-like G protein-coupled receptor (5-HT-GPCR).
XX
KW Human; serotonin-like G protein-coupled receptor; 5-HT-GPCR;
KW peripheral nervous system; central nervous system; PNS; CNS;
KW brain injury; mood disorder; anxiety disorder; sleep disorder;
KW neurogenic; myopathic disorder; neurodegenerative disorder;
KW tranquillizer; nootropic; neuroprotective; antiparkinsonian;
KW analgesic; cerebroprotective.
XX
OS Homo sapiens.
XX
PN WO200170967-A2.
XX
PD 27-SEP-2001.
XX
PF 19-MAR-2001; 2001WO-EP03117.
XX
PR 20-MAR-2000; 2000US-190104P.
PR 12-JUN-2000; 2000US-210975P.
PR 14-DEC-2000; 2000US-255110P.
XX
PA (FARB) BAYER AG.
XX
PI Ramakrishnan S;
XX
DR WPI; 2001-611498/70.
DR N-PSDB; AAS12583.
XX
PT New polynucleotide encoding a polypeptide which regulates, prevents and
PT treats diseases of the peripheral or central nervous system including
PT Alzheimer's Disease, comprises the human serotonin-like G
PT protein-coupled receptor polynucleotide -
XX
PS Claim 1; Fig 2; 88pp; English.
XX
CC The present invention relates to the isolation of a novel DNA sequence
CC encoding a human serotonin-like G protein-coupled receptor (5-HT-GPCR)
CC polypeptide. The sequences of the invention are useful for screening for
CC agents which decrease the activity of 5-HT-GPCR or for identifying agents
CC which regulate the activity of 5-HT-GPCR. A reagent that modulates the
CC activity of 5-HT-GPCR is useful for detecting 5-HT-GPCR in a biological
CC sample and for reducing the activity of 5-HT-GPCR in a cell. A
CC pharmaceutical composition comprising such a reagent is useful for
CC preventing or ameliorating disorders of the peripheral or central nervous
CC system, preferably primary or secondary disorders after brain injury,
CC mood disorders, anxiety disorders, disorders of thought and volition,
CC sleep disorders, diseases of the motor unit such as neurogenic and
CC myopathic disorders, neurodegenerative disorders (e.g. Alzheimer's
CC disease, Parkinson's disease) or disorders leading to peripheral and
CC chronic pain. The coding sequence of 5-HT-GPCR polynucleotide is useful
CC for generating antisense oligonucleotides or ribozymes which specifically
CC bind to mRNA transcribed from the 5-HT-GPCR polynucleotide. These
CC antisense oligonucleotides are useful for modulating 5-HT-GPCR gene

```
CC   expression. Polynucleotide sequences encoding for 5-HT-GPCR may be used
CC   in gene therapy. The present sequence represents the novel human
CC   5-HT-GPCR polypeptide of the invention.
XX
SQ   Sequence   451 AA;

  Query Match            100.0%;  Score 2318;  DB 22;  Length 451;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-228;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
           |||||||||||||||||||||||||||||||
Db     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451


RESULT 3
AAU08710
ID   AAU08710 standard; Protein; 451 AA.
XX
AC   AAU08710;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human histamine H2-like G protein-coupled receptor polypeptide.
XX
KW   Histamine H2-like G protein-coupled receptor; histamine H2-like GPCR;
KW   digestive system; immune system; respiratory system; reproductive system;
KW   urinary system; peripheral nervous system; central nervous system; human;
KW   brain injury; mood disorder; anxiety; thought disorder; sleep disorder;
KW   motor unit disease; neurogenic disorder; myopathic disorder; neuroleptic;
KW   neurodegenerative disorder; psychotic disorder; cerebrovascular disorder;
KW   Alzheimer's disease; Parkinson's disease; muscle spasm; peripheral pain;
KW   chronic pain; tranquiliser; nootropic; neuroprotective; antiparkinsonian;
KW   analgesic; antidepressant; antiallergic; antiinflammatory; gene therapy;
KW   cerebroprotective; vulnerary.
XX
OS   Homo sapiens.
XX
PN   WO200170812-A2.
XX
PD   27-SEP-2001.
XX
PF   20-MAR-2001; 2001WO-EP03162.
XX
PR   20-MAR-2000; 2000US-190554P.
PR   12-JUN-2000; 2000US-210734P.
PR   14-DEC-2000; 2000US-255147P.
XX
PA   (FARB ) BAYER AG.
XX
PI   Ramakrishnan S;
XX
DR   WPI; 2001-611486/70.
DR   N-PSDB; AAS14725.
XX
PT   New human histamine H2-like G protein-coupled receptor polynucleotide
PT   and polypeptide which can be regulated for preventing, treating
PT   diseases of digestive, immune, respiratory, reproductive or central
PT   nervous system  -
XX
PS   Claim 1; Fig 2; 93pp; English.
XX
CC   The invention relates to a human histamine H2-like G protein-coupled
CC   receptor (GPCR) polypeptides and the nucleic acids encoding them. The
CC   sequences are useful for screening for agents which regulate the activity
CC   of histamine H2-like GPCR by contacting test compounds with the
CC   polypeptide and monitoring the activity. The polypeptides and their
CC   associated DNA sequences are useful for modulating the activity of a
CC   histamine H2-like GPCR in a disease of the digestive, immune,
CC   respiratory, reproductive, urinary, peripheral or central nervous system,
CC   especially a primary or secondary disorder after brain injury, disorder
CC   of mood, an anxiety disorder, a disorder of thought, a disorder of sleep,
CC   a disease of the motor unit, a neurogenic and myopathic disorder, a
CC   neurodegenerative disorder, a psychotic disorder, a cerebrovascular
CC   disorder, Alzheimer's disease, Parkinson's disease, muscle spasms or a
CC   disorder leading to peripheral and chronic pain. This sequence represents
CC   a human histamine H2-like GPCR polypeptide.
XX
SQ   Sequence   451 AA;

  Query Match            100.0%;  Score 2318;  DB 22;  Length 451;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-228;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
           |||||||||||||||||||||||||||||||
Db     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451


RESULT 4
```

```
AAU04367
ID   AAU04367 standard; Protein; 451 AA.
XX
AC   AAU04367;
XX
DT   23-OCT-2001  (first entry)
XX
DE   Human G-protein coupled receptor, hRUP13.
XX
KW   Human; G-protein coupled receptor; GPCR; hRUP13; agonist;
KW   inverse agonist; lung cancer.
XX
OS   Homo sapiens.
XX
PN   WO200136471-A2.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31509.
XX
PR   17-NOV-1999;   99US-0166088.
PR   17-NOV-1999;   99US-0166099.
PR   17-NOV-1999;   99US-0166369.
PR   23-DEC-1999;   99US-0171900.
PR   23-DEC-1999;   99US-0171901.
PR   23-DEC-1999;   99US-0171902.
PR   11-FEB-2000; 2000US-0181749.
PR   14-MAR-2000; 2000US-0189258.
PR   14-MAR-2000; 2000US-0189259.
PR   10-APR-2000; 2000US-0195898.
PR   10-APR-2000; 2000US-0195899.
PR   10-APR-2000; 2000US-0196078.
PR   28-APR-2000; 2000US-0200419.
PR   12-MAY-2000; 2000US-0203630.
PR   12-JUN-2000; 2000US-0210741.
PR   12-JUN-2000; 2000US-0210982.
PR   21-AUG-2000; 2000US-0226760.
PR   26-SEP-2000; 2000US-0235418.
PR   26-SEP-2000; 2000US-0235779.
PR   20-OCT-2000; 2000US-0242332.
PR   20-OCT-2000; 2000US-0242343.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Chen R,  Dang HT,  Lowitz KP;
XX
DR   WPI; 2001-355616/37.
DR   N-PSDB; AAS07940.
XX
PT   Endogenous and non-endogenous versions of human G-protein coupled
PT   receptors for direct identification of candidate compounds as agonists,
PT   inverse agonists or partial agonists for use as therapeutic agents -
XX
PS   Claim 21; Page 98-100; 160pp; English.
XX
CC   The sequence represents a human G-protein coupled receptor (GPCR),
CC   hRUP13. The endogenous and non-endogenous, constitutively activated
CC   versions of human G-protein coupled receptors (GPCR), are useful for
CC   direct identification of candidate compounds as receptor agonists,
CC   inverse agonists or partial agonists having applicability as therapeutic
CC   agents for treating diseases related to GPCR, e.g. lung cancer.
CC   Non-endogenous version of human GPCRs are also utilized in research
CC   settings and in vitro and in vivo system, incorporating GPCRs can be
CC   utilised to elucidate and understand the roles these receptors
CC   play in the human condition, both normal and diseased.
XX
SQ   Sequence   451 AA;

  Query Match             100.0%;  Score 2318;  DB 22;  Length 451;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-228;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
                 |||||||||||||||||||||||||||||
Db     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451


RESULT 5
AAG64123
ID   AAG64123 standard; Protein; 451 AA.
XX
AC   AAG64123;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Human G protein-coupled receptor GPRv47.
XX
KW   Human; guanosine triphosphate binding protein-coupled receptor;
KW   G protein-coupled receptor; GPRv8; GPRv12; GPRv16; GPRv21; GPRv40;
KW   GPRv47; GPRv51; GPRv71; GPRv72; cancer; liver cirrhosis;
KW   Alzheimer's disease; cytostatic; hepatotropic; nootropic;
KW   neuroprotective; gene therapy; peptide therapy.
XX
OS   Homo sapiens.
XX
PN   WO200148188-A1.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-JP09408.
XX
PR   28-DEC-1999;   99JP-0375152.
PR   31-MAR-2000; 2000JP-0101339.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Matsumoto S,  Oda T,  Saito Y,  Morikawa N,  Yoshida K,  Suwa M;
PI   Sugiyama T,  Kishimoto T,  Kanzaki K,  Yasuda S,  Inoue Y;
XX
DR   WPI; 2001-425662/45.
DR   N-PSDB; AAH73514.
XX
PT   New DNA encoding guanosine triphosphate binding protein coupled
PT   receptors and their expression products for screening potential
PT   anticancer and nootropic drugs and in diagnosis of these diseases -
XX
PS   Example 1; Page 126-129; 170pp; Japanese.
XX
```

```
CC   The invention relates to nine human guanosine triphosphate binding
CC   protein (G protein)-coupled receptors designated GPRv8, GPRv12, GPRv16,
CC   GPRv21, GPRv40, GPRv47, GPRv51, GPRv71 and GPRv72, and to the
CC   genes encoding them. These genes and proteins and antibodies against
CC   the protein are useful in the treatment, prevention, diagnosis and
CC   investigation of diseases associated with G protein-coupled receptors,
CC   including cancer, cirrhosis of the liver and Alzheimer's disease.
CC   The present sequence is a G protein-coupled receptor of the invention.
XX
SQ   Sequence   451 AA;

  Query Match             100.0%;  Score 2318;  DB 22;  Length 451;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-228;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
           |||||||||||||||||||||||||||||||
Db     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451


RESULT 6
ABB81539
ID   ABB81539 standard; Protein; 451 AA.
XX
AC   ABB81539;
XX
DT   03-SEP-2002  (first entry)
XX
DE   Human G-protein coupled receptor protein SEQ ID NO:2.
XX
KW   Human; G-protein coupled receptor; GPCR; receptor; chromosome 1.
XX
OS   Homo sapiens.
XX
PN   WO200250272-A2.
XX
PD   27-JUN-2002.
XX
PF   19-DEC-2001; 2001WO-US48547.
XX
PR   20-DEC-2000; 2000US-0740033.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Gan W,  Di Francesco V,  Beasley EM;
XX
DR   WPI; 2002-500628/53.
DR   N-PSDB; ABN89471, ABN89472.
XX
PT   Novel isolated G-protein coupled receptor peptide useful for treating
PT   disorder characterized by absence of, inappropriate or unwanted
PT   expression of the receptor protein, and as immunogens to raise
PT   antibodies -
XX
PS   Claim 1; Fig 2A; 72pp; English.
XX
CC   The present sequence represents a human G-protein coupled receptor (GPCR)
CC   protein (I). (I) can be used in gene therapy. (I) can be used for
CC   identifying a modulator of (I) by contacting (I) with an agent and
CC   determining if the agent has modulated the function or activity of (I).
CC   (I) is also useful for identifying an agent that binds to (I), by
CC   contacting (I) with an agent and assaying the contacted mixture to
CC   determine whether a complex is formed with the agent bound to (I).
CC   (I) can be used in the treatment of a disorder characterised by the
CC   absence of, inappropriate or unwanted expression of the receptor
CC   protein. (I) is located to human chromosome 1.
XX
SQ   Sequence   451 AA;

  Query Match             100.0%;  Score 2318;  DB 23;  Length 451;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-228;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
           |||||||||||||||||||||||||||||||
Db     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451


RESULT 7
AAU79732
ID   AAU79732 standard; Protein; 451 AA.
XX
AC   AAU79732;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Novel human GPCR, PFI-018.
XX
KW   Human; G-protein coupled receptor; GPCR; obesity; diabetes; inflammation;
KW   metabolic disease; neurological disease; psychotherapeutic; cancer;
```

```
KW   urogenital disease; reproductive and sexual disease; tissue repair;
KW   dermatology; skin pigmentation; photoageing; frailty; osteoporosis;
KW   cardiovascular disease; gastrointestinal disease; infection; allergy;
KW   respiratory disease; sensory organ disorder; sleep disorder; hair loss;
KW   PFI-018; receptor.
XX
OS   Homo sapiens.
XX
PN   EP1158002-A1.
XX
PD   28-NOV-2001.
XX
PF   16-MAY-2001; 2001EP-0304335.
XX
PR   19-MAY-2000; 2000GB-0012248.
PR   23-MAR-2001; 2001GB-0007394.
XX
PA   (PFIZ ) PFIZER LTD.
PA   (PFIZ ) PFIZER INC.
XX
PI   Lewis ME,  Robas NM;
XX
DR   WPI; 2002-306187/35.
DR   N-PSDB; ABK48871.
XX
PT   An isolated polynucleotide (I) encoding a G-protein coupled receptor is
PT   useful in treating disorders e.g. diabetes -
XX
PS   Claim 22; Fig 2; 50pp; English.
XX
CC   The present invention relates to the isolation of a novel human
CC   G-protein coupled receptor (GPCR), PFI-018, and the polynucleotide
CC   sequence encoding it. Therapeutically the polynucleotide encoding
CC   PFI-018 may be useful in treating obesity, diabetes, metabolic diseases,
CC   neurological diseases, psychotherapeutics, urogenital diseases,
CC   reproductive and sexual diseases, inflammations, cancers, tissue repair,
CC   dermatology, skin pigmentation, photoageing, frailty, osteoporosis,
CC   cardiovascular disease, gastrointestinal disease, infections, allergies,
CC   respiratory disease, sensory organ disorders, sleep disorders and hair
CC   loss. The present sequence represents human GPCR, PFI-018.
XX
SQ   Sequence   451 AA;

  Query Match             100.0%;  Score 2318;  DB 23;  Length 451;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-228;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
           |||||||||||||||||||||||||||||||
Db     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451


RESULT 8
AAM50637
ID   AAM50637 standard; Protein; 451 AA.
XX
AC   AAM50637;
XX
DT   04-APR-2002  (first entry)
XX
DE   Human G protein coupled receptor IGPcR27.
XX
KW   IGPcR27; G-protein coupled receptor; receptor; human;
KW   Tourette syndrome; Parkinson's disease; Huntington's disease;
KW   tic; spasm; tremor; dyskinesia; dystonia; pain; analgesic;
KW   neuroleptic; tranquillizer; antiparkinsonian; neuroprotective;
KW   nootropic; anticonvulsant; metabolic; anorectic; anabolic;
KW   antiinflammatory; antidiarrheic; osteopathic; antibacterial;
KW   fungicide; protozoacide; virucide; antiasthmatic; antiallergic;
KW   antiarthritic; immunosuppressive; nephrotropic; diagnosis;
KW   central nervous system; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          44..66
FT                   /label= Transmembrane_domain-1
FT   Domain          78..98
FT                   /label= Transmembrane_domain-2
FT   Domain          116..137
FT                   /label= Transmembrane_domain-3
FT   Domain          160..177
FT                   /label= Transmembrane_domain-4
FT   Domain          210..230
FT                   /label= Transmembrane_domain-5
FT   Domain          291..309
FT                   /label= Transmembrane_domain-6
FT   Domain          324..346
FT                   /label= Transmembrane_domain-7
XX
PN   WO200202600-A2.
XX
PD   10-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP07534.
XX
PR   30-JUN-2000; 2000US-215881P.
XX
PA   (INGE-) INGENIUM PHARM AG.
XX
PI   Wattler F,  Wattler S,  Trommler P,  Nehls MC;
XX
DR   WPI; 2002-140081/18.
DR   N-PSDB; ABA91237.
XX
PT   New human or mouse G protein-coupled receptor protein, IGPcR27, useful
PT   for diagnosis, prevention, amelioration or treatment of central nervous
PT   system disorders such as Tourette's syndrome, Parkinson's disease and
PT   pain  -
XX
PS   Claim 8; Fig 2; 71pp; English.
XX
CC   The present sequence is that of human IGPcR27, a novel G protein
CC   coupled receptor expressed exclusively in the brain.  The invention
CC   provides human and mouse IGPcR27 polynucleotides and polypeptides,
CC   vectors, host cell expression systems, and transgenic and knockout
CC   animals.  Vectors and host cells are used in gene therapy for
CC   prevention, amelioration or treatment of diseases characterized by
```

```
CC   aberrant expression or activity of IGPcR27, where the disease is
CC   associated with signal processing in the central nervous system (CNS)
CC   such as movement dysfunctions, disorders or diseases, tics, tremor,
CC   Tourette's syndrome, Parkinson's disease, Huntington's disease,
CC   dyskinesias, dystonia, pain and spasms (claimed). Knockout,
CC   mutated and transgenic animals are useful for the dissection of the
CC   molecular mechanisms of the IGPcR27 pathway, for the identification
CC   and cloning of genes able to modify, reduce or inhibit the phenotype
CC   associated with IGPcR27 activity or deficiency, for the
CC   identification of gene and protein diagnostic markers for diseases,
CC   and for the identification and testing of compounds useful in the
CC   prevention, amelioration or treatment of diseases associated with
CC   IGPcR27 activity or deficiency (claimed). IGPcR27 proteins,
CC   nucleic acids, antibodies, agonists and antagonists are useful for
CC   the diagnosis and treatment of CNS disorders, for the detection of
CC   mutant forms of IGPcR27 or inappropriately expressed forms of
CC   IGPcR27, and for drug screening. IGPcR27 protein is useful in
CC   diagnosis, prevention, amelioration or treatment of diseases
CC   associated with signal processing in CNS, schizophrenia, episodic
CC   paroxysmal anxiety (EPA) disorders such as obsessive compulsive
CC   disorder (COD), multiple sclerosis, Alzheimer's disease/dementia,
CC   anorexia, kidney diseases such as renal failure, obesity,
CC   gastrointestinal disorders such as irritable bowel syndrome (IBS),
CC   diarrhoea, motility disorders and conditions of delayed gastric
CC   emptying, osteoporosis, infections such as bacterial, fungal,
CC   protozoal and viral infections, asthma, allergy, arthritis,
CC   sepsis and gynaecological disorders. The proteins and nucleic
CC   acids are useful for the identification of compounds effective in
CC   the treatment of disorders based on the aberrant expression or
CC   activity of IGPcR27. IGPcR27 is also useful for the generation of
CC   antibodies, in identification of other cellular gene products
CC   involved in regulating IGPcR27, and as a pharmaceutical reagent.
XX
SQ   Sequence   451 AA;

  Query Match            100.0%;  Score 2318;  DB 23;  Length 451;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-228;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
           |||||||||||||||||||||||||||||||
Db     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451

RESULT 9
AAO14014
ID   AAO14014 standard; Protein; 451 AA.
XX
AC   AAO14014;
XX
DT   05-APR-2002  (first entry)
XX
DE   Human G protein-coupled receptor Con-218.
XX
KW   Con-218; G protein-coupled receptor; GPCR; anti-HIV; antiparkinsonian;
KW   neuroprotective; cytostatic; tranquilliser; neuroleptic; antimanic;
KW   antidepressant; immunosuppressive; antimigraine; nootropic; cardiant;
KW   antiarteriosclerotic; antidiabetic; thrombolytic; antipsoriatic;
KW   vasotropic; anticonvulsant; antithyroid; antiinflammatory; nephrotropic;
KW   hypotensive; antirheumatic; antiarthritic; cerebroprotective; virucide;
KW   antiinfertility; gene therapy; thyroid disorder; renal failure;
KW   inflammatory conditions; cell differentiation; homeostasis; CNS disorder;
KW   rheumatoid arthritis; autoimmune disorder; movement disorder; stroke;
KW   psychotic disorder; neurological disorder; dyskinesia; infection;
KW   attention disorder; degenerative disorder; metabolic; cardiovascular;
KW   cancer; hyperproliferative disorder; psoriasis; hormonal disorder;
KW   sexual dysfunction; schizophrenia; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          44..66
FT                   /label= Transmembrane_domain
FT   Domain          77..99
FT                   /label= Transmembrane_domain
FT   Domain          118..136
FT                   /label= Transmembrane_domain
FT   Domain          156..178
FT                   /label= Transmembrane_domain
FT   Domain          207..229
FT                   /label= Transmembrane_domain
FT   Domain          285..307
FT                   /label= Transmembrane_domain
FT   Domain          326..344
FT                   /label= Transmembrane_domain
XX
PN   WO200181576-A2.
XX
PD   01-NOV-2001.
XX
PF   19-APR-2001; 2001WO-US12690.
XX
PR   19-APR-2000; 2000US-198600P.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Lind P,  Berthold M;
XX
DR   WPI; 2002-041403/05.
DR   N-PSDB; AAK98269.
XX
PT   Novel gene encoding G protein-coupled receptor termed Con-218, useful
PT   for identifying compounds that are useful for treating inflammatory
PT   conditions, central nervous system disorders and cardiovascular
PT   disorders -
XX
PS   Claim 31; Page 73; 126pp; English.
XX
CC   The sequence represents Con-218, the novel human G protein-coupled
CC   receptor (GPCR) of the invention. The polypeptide of the invention has
CC   anti-HIV, antiparkinsonian, neuroprotective, cytostatic, tranquilliser,
CC   neuroleptic, antimanic, antidepressant, immunosuppressive, antimigraine,
CC   nootropic, antiarteriosclerotic, cardiant, antidiabetic, thrombolytic,
CC   antipsoriatic, vasotropic, anticonvulsant, antithyroid, antiinflammatory,
CC   nephrotropic, hypotensive, antirheumatic, antiarthritic,
CC   cerebroprotective, antiinfertility, and virucide activity. The Con-218
CC   works as a modulator of neuropeptide binding or signalling, in a vaccine
CC   or gene therapy. The polypeptide and polynucleotide are useful for the
```

```
CC   treatment of thyroid disorders; renal failure; inflammatory conditions;
CC   cell differentiation and homeostasis diseases; rheumatoid arthritis;
CC   autoimmune disorders; movement disorders; CNS disorders; stroke;
CC   psychotic and neurological disorders; dyskinesias; attention disorders;
CC   degenerative disorders; infections; metabolic and cardiovascular
CC   diseases and disorders; proliferative diseases and cancers;
CC   hyperproliferative disorders; psoriasis; hormonal disorders; and sexual
CC   dysfunction. The polypeptide is also useful for inducing an immune
CC   response in a mammal against itself, and for identifying a compound that
CC   binds Con-281. Con-218 molecules are useful for identifying a compound
CC   which modulates the activity of Con-218 polypeptides. The polypeptide is
CC   also useful for identifying compounds useful for the treatment of
CC   schizophrenia. Furthermore, the polypeptide is useful for purifying a G
CC   protein from a sample containing G-protein.
XX
SQ   Sequence   451 AA;

  Query Match            100.0%;  Score 2318;  DB 23;  Length 451;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-228;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
           |||||||||||||||||||||||||||||||
Db     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451


RESULT 10
AAU74909
ID   AAU74909 standard; Protein; 451 AA.
XX
AC   AAU74909;
XX
DT   09-APR-2002  (first entry)
XX
DE   Amino acid sequence of human G-protein coupled receptor TGR213 protein.
XX
KW   Human; G-protein coupled; receptor; GPCR; TGR213; kidney disease;
KW   signal transduction modulator; cerebral cavernous malformation;
KW   hyperlipidemia; obesity; dyslexia; cardiac myxoma; renal failure;
KW   nephritis; hypertension; liver disease; cirrhosis; blood disorder;
KW   spleen-associated disorder; immune disorder.
XX
OS   Homo sapiens.
XX
PN   WO200200719-A2.
XX
PD   03-JAN-2002.
XX
PF   25-JUN-2001; 2001WO-US20363.
XX
PR   23-JUN-2000; 2000US-213461P.
XX
PA   (TULA-) TULARIK INC.
XX
PI   Lin DC,  Zhao J,  Chen J,  Cutler G;
XX
DR   WPI; 2002-147880/19.
DR   N-PSDB; ABK12962.
XX
PT   New G-protein coupled receptor polypeptides, useful for identifying
PT   modulators of signal transduction for treating kidney disease,
PT   hyperlipidemia, obesity, dyslexia and cardiac myxoma  -
XX
PS   Claim 26; Page 65; 78pp; English.
XX
CC   The present invention relates to a new G-protein coupled receptor (GPCR)
CC   polypeptide comprising greater than 70% amino acid sequence identity to
CC   the amino acid sequence of human GPCRs TGR62, TGR21, TGR130.1, TGR130.2,
CC   human TGR213 or TGR92, 80% amino acid sequence identity to mouse TGR18
CC   or 90% amino acid sequence identity to human novel edg receptor protein,
CC   as defined in the specification. The GPCR covalently linked to a solid
CC   phase is useful for identifying a compound that modulates signal
CC   transduction. The identified compounds are useful for treating
CC   kidney disease, cerebral cavernous malformations, hyperlipidemia,
CC   obesity, dyslexia and cardiac myxoma. The molecules of the invention are
CC   useful for diagnosing disorders or conditions such as kidney-related
CC   conditions or diseases such as renal failure, nephritis, nephrotic
CC   syndrome, asymptomatic urinary abnormalities, renal tubule defects,
CC   hypertension and nephrolithiasis, liver-related disease or condition
CC   e.g. cirrhosis, infiltrations, lesions, functional disorders and jaundice
CC   and spleen-associated disorders or conditions e.g. splenic enlargement,
CC   immune disorders, blood disorders and others. Modulation of the
CC   polypeptide of the invention is useful to treat or prevent any of the
CC   above conditions or diseases. The present amino acid sequence represents
CC   the human GPCR TGR213 protein of the invention. This sequence is one of
CC   seven novel G protein coupled receptors of the invention (AAU74904-
CC   AAU74911).
XX
SQ   Sequence   451 AA;

  Query Match            100.0%;  Score 2318;  DB 23;  Length 451;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-228;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
```

```
Qy    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
          |||||||||||||||||||||||||||||||
Db    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451


RESULT 11
AAU11895
ID   AAU11895 standard; Protein; 451 AA.
XX
AC   AAU11895;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human novel G protein-coupled receptor, GPCR8a.
XX
KW   Human; GPCR8a; G protein coupled receptor; developmental disease;
KW   immune disease; taste and scent detectability disorder;
KW   Burkitt's lymphoma; signal transduction pathway disorder;
KW   retinal disease; cell growth rate disorder; feeding disorder;
KW   control of feeding; obesity; starvation; fungal infection;
KW   noninsulin-dependent diabetes mellitus; bacterial infection;
KW   protozoal infection; viral infection; pain; cancer; anorexia; bulimia;
KW   asthma; Parkinson's disease; acute heart failure; hypotension;
KW   hypertension; urinary retention; osteoporosis; Crohn's disease;
KW   multiple sclerosis; Albright Hereditary Osteodystrophy; angina pectoris;
KW   myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;
KW   manic depression; delirium; dementia; severe metal retardation;
KW   autosomal dominant (2) Acrocallosal syndrome; dyskinesia;
KW   Huntington's disease; Gilles de la Tourette syndrome;
KW   haematopoietic disorder; cell signal processing disorder;
KW   metabolic pathway modulation disorder.
XX
OS   Homo sapiens.
XX
PN   WO200190187-A2.
XX
PD   29-NOV-2001.
XX
PF   24-MAY-2001; 2001WO-US17114.
XX
PR   24-MAY-2000; 2000US-206757P.
PR   25-MAY-2000; 2000US-207020P.
PR   28-JUN-2000; 2000US-214372P.
PR   19-JUL-2000; 2000US-219786P.
PR   25-JUL-2000; 2000US-220593P.
PR   10-OCT-2000; 2000US-239542P.
PR   18-DEC-2000; 2000US-256402P.
PR   26-FEB-2001; 2001US-271645P.
PR   09-MAR-2001; 2001US-274809P.
PR   13-MAR-2001; 2001US-275590P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Padigaru M,  Spytek KA,  Majumder K,  Tchernev VT,  Grosse WM;
PI   Szerkeres ED,  Alsobrook JP,  Burgess CE,  Shimkets RA,  Taupier RJ;
PI   Casman SJ,  Gangolli E,  MacDougall JR,  Stone DJ,  Smithson G;
XX
DR   WPI; 2002-106188/14.
DR   N-PSDB; AAS19412.
XX
PT   New polypeptide, useful for treating pain, cancer, urinary retention,
PT   osteoporosis, Crohn's disease, dyskinesias, acute heart failure,
PT   dementia, asthma and ulcer, comprises the isolated G-coupled
PT   protein-receptor GPCRX  -
XX
PS   Claim 1; Page 35; 154pp; English.
XX
CC   The invention relates to an isolated G protein-coupled receptor related
CC   polypeptide (GPCR1a, 1b, 2, 3a, 3b, 4, 5, 6a, 6b, 7, 8a, 8b, and GPCR9)
CC   and the polynucleotides encoding them. The GPCR, its encoding
CC   polynucleotide and an anti-GPCR antibody are useful for diagnosing
CC   treating or preventing a GPCR-associated disorder in a human e.g.
CC   developmental diseases, immune diseases, taste and scent detectability
CC   disorder, Burkitt's lymphoma, signal transduction pathway disorders,
CC   retinal diseases including those involving photoreception, cell growth
CC   rate disorders, feeding disorders, control of feeding, potential
CC   obesity due to over-eating, potential disorders due to starvation,
CC   noninsulin-dependent diabetes mellitus, bacterial, fungal, protozoal
CC   and viral infections, pain, cancer, anorexia, bulimia, asthma,
CC   Parkinson's disease, acute heart failure, hypotension, hypertension,
CC   urinary retention, osteoporosis, Crohn's disease, multiple sclerosis,
CC   Albright Hereditary Osteodystrophy, angina pectoris, myocardial
CC   infarction, ulcers, allergies, benign prostatic hypertrophy, manic
CC   depression, delirium, dementia, severe metal retardation, autosomal
CC   dominant (2) Acrocallosal syndrome and dyskinesias, such as
CC   Huntington's disease or Gilles de la Tourette syndrome, haematopoietic
CC   disorders, and disorders related to cell signal processing and
CC   metabolic pathway modulation. The present sequence represents GPCR8a.
XX
SQ   Sequence   451 AA;

  Query Match             100.0%;  Score 2318;  DB 23;  Length 451;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-228;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy     61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy    121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
          |||||||||||||||||||||||||||||||
Db    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451


RESULT 12
AAE17073
ID   AAE17073 standard; Protein; 451 AA.
XX
AC   AAE17073;
XX
DT   18-APR-2002  (first entry)
XX
DE   Human G-protein coupled receptor (GPCRx2) protein.
XX
KW   Human; G-protein coupled receptor; GPCRx2; cerebroprotective; vomiting;
KW   receptor-mediated disorder; therapy; urinary retention; allergy; obesity;
KW   osteoporosis; angina pectoris; restenosis; atherosclerosis; hypotension;
KW   anorexia; tumour; migraine; acute heart failure; ulcer; antiinflammatory;
```

```
KW   stroke; hypertension; neuronal disorder; myocardial infarction psychotic;
KW   depression; mental retardation; neurodegenerative disease; antibacterial;
KW   Alzheimer's disease; dementia; ischaemia; Parkinson's disease; antiviral;
KW   Huntington's disease; anxiety; antifungal; immunosuppressive; cytostatic;
KW   vulnerary; analgesic; anorectic; anabolic; diuretic; cardiant; nootropic;
KW   antiemetic; vasotropic; diabetes; cancer; tranquillizer; neuroleptic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          45..68
FT                   /note= "transmembrane domain"
FT   Domain          78..99
FT                   /note= "transmembrane domain"
FT   Domain          117..138
FT                   /note= "transmembrane domain"
FT   Domain          154..179
FT                   /note= "transmembrane domain"
FT   Domain          207..228
FT                   /note= "transmembrane domain"
FT   Domain          287..305
FT                   /note= "transmembrane domain"
FT   Domain          325..344
FT                   /note= "transmembrane domain"
XX
PN   WO200198330-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-BE00104.
XX
PR   20-JUN-2000; 2000US-212913P.
PR   11-JUL-2000; 2000US-217494P.
PR   26-JAN-2001; 2001EP-0870015.
PR   12-FEB-2001; 2001EP-0870024.
XX
PA   (EURO-) EUROSCREEN SA.
XX
PI   Lannoy V,  Brezillon S,  Detheux M,  Parmentier M,  Govarts C;
XX
DR   WPI; 2002-130789/17.
DR   N-PSDB; AAD27493.
XX
PT   New G-protein coupled receptor, useful in the manufacture of
PT   medicaments for treating receptor mediated disorders e.g. acute heart
PT   failure and Alzheimer's disease -
XX
PS   Disclosure; Page 21; 46pp; English.
XX
CC   The present invention relates to a G-protein coupled receptor (GPCR) and
CC   nucleotide encoding it. GPCR are useful in the manufacture of a
CC   medicament for the prevention and/or treatment of receptor-mediated
CC   disorders e.g. viral infections, virus and bacterial diseases, diseases
CC   and disorders involving disturbances of cell migration, diseases or
CC   perturbations of immune system including cancers, development of tumours
CC   and tumour metastasis, inflammatory and neoplastic processes; bacterial
CC   and fungal infections, in wound and bone healing, dysfunction of
CC   regulatory growth functions; pains, diabetes, obesity, anorexia, bulimia,
CC   urinary retention, osteoporosis, angina pectoris, atherosclerosis,
CC   restenosis, diseases involving excessive or reduced proliferation or loss
CC   of smooth muscle cells, aneurysm, stroke, ischaemia, ulcers, allergies,
CC   benign prostatic hypertrophy, migraine, vomiting; blood circulating
CC   affections including acute heart failure, hypotension, hypertension and
CC   myocardial infarction psychotic; neuronal disorders such as anxiety,
CC   schizophrenia, maniac depression, depression, delirium, dementia, severe
CC   mental retardation; degenerative diseases; neurodegenerative diseases
CC   such as Alzheimer's disease, Parkinson's disease; and dyskinesias e.g.
CC   Huntington's disease or Gilles de la Tourette's syndrome and other
CC   related diseases. The present sequence is GPCRx2 protein.
XX
SQ   Sequence   451 AA;

 Query Match             100.0%;  Score 2318;  DB 23;  Length 451;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-228;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
           |||||||||||||||||||||||||||||||
Db     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451


RESULT 13
ABB05445
ID   ABB05445 standard; Protein; 451 AA.
XX
AC   ABB05445;
XX
DT   16-APR-2002  (first entry)
XX
DE   Human G-protein coupled receptor GPRMDA12 protein SEQ ID NO:2.
XX
KW   Human; G-protein coupled receptor; GPRMDA12; receptor; antibacterial;
KW   fungicide; protozoacide; virucide; anti-HIV; analgesic; cytostatic; pain;
KW   antidiabetic; anorectic; antiasthmatic; antiparkinsonian; cardiant;
KW   hypertensive; hypotensive; osteopathic; antianginal; cerebroprotective;
KW   antiulcer; antiallergic; antimigraine; antiemetic; tranquilliser; cancer;
KW   neuroleptic; antimanic; antidepressant; anticonvulsant; nootropic; ulcer;
KW   vaccine; gene; infection; diabetes; obesity; anorexia; bulimia; asthma;
KW   Parkinson's disease; acute heart failure; hypotension; hypertension;
KW   urinary retention; osteoporosis; angina pectoris; stroke; schizophrenia;
KW   myocardial infarction; allergy; benign prostatic hypertrophy; migraine;
KW   vomiting; psychotic; neurological disorder; anxiety; manic depression;
KW   depression; delirium; dementia; severe mental retardation; dyskinesia;
KW   Huntington's disease; Gilles dela Tourette's syndrome.
XX
OS   Homo sapiens.
XX
PN   WO200198490-A1.
XX
PD   27-DEC-2001.
XX
PF   15-JUN-2001; 2001WO-EP06729.
XX
PR   21-JUN-2000; 2000EP-0113239.
XX
PA   (MERE ) MERCK PATENT GMBH.
XX
PI   Kluxen F,  Duecker K;
```

```
XX
DR   WPI; 2002-139788/18.
DR   N-PSDB; ABA93134.
XX
PT   Novel G-protein coupled receptor polypeptides and polynucleotides
PT   useful in screening assays and for treating diabetes, cancer,
PT   osteoporosis, asthma, neurological disorders and identifying modulators
PT   useful in therapy -
XX
PS   Claim 1; Page 40-41; 44pp; English.
XX
CC   The present sequence represents a human G-protein coupled receptor
CC   designated GPRMDA12 (I). (I) has antibacterial, fungicide, protozoacide,
CC   virucide, anti-HIV, analgesic, cytostatic, antidiabetic, anorectic,
CC   antiasthmatic, antiparkinsonian, cardiant, hypertensive, hypotensive,
CC   osteopathic, antianginal, cerebroprotective, antiulcer, antiallergic,
CC   antimigraine, antiemetic, tranquilliser, neuroleptic, antidepressant,
CC   antimanic, anticonvulsant and nootropic activities, and it can be used
CC   in vaccine production. (I) can be used in screening assays to identify
CC   compounds that stimulate or inhibit the function or level of (I). (I)
CC   and the polynucleotide (II) encoding it can be used in treating and
CC   preventing human diseases including infections, particularly infections
CC   caused by HIV-1 or HIV-2, pain, cancers, diabetes, obesity, anorexia,
CC   bulimia, asthma, Parkinson's disease, acute heart failure, hypotension,
CC   hypertension, urinary retention, osteoporosis, angina pectoris, stroke,
CC   myocardial infarction, ulcer, allergy, benign prostatic hypertrophy,
CC   migraine, vomiting, psychotic and neurological disorders, including
CC   anxiety, schizophrenia, manic depression, depression, delirium, dementia,
CC   severe mental retardation, dyskinesias such as Huntington's disease or
CC   Gilles dela Tourette's syndrome.
XX
SQ   Sequence   451 AA;

  Query Match             100.0%;  Score 2318;  DB 23;  Length 451;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-228;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
           |||||||||||||||||||||||||||||||
Db     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451


RESULT 14
AAU04386
ID   AAU04386 standard; Protein; 869 AA.
XX
AC   AAU04386;
XX
DT   23-OCT-2001  (first entry)
XX
DE   GPCR-Gs fusion protein, hRUP13-GS.
XX
KW   G-protein coupled receptor; GPCR; hRUP13-Gs; agonist;
KW   inverse agonist; lung cancer.
XX
OS   Chimeric - Homo sapiens.
OS   Chimeric - Rattus sp.
XX
PN   WO200136471-A2.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31509.
XX
PR   17-NOV-1999;   99US-0166088.
PR   17-NOV-1999;   99US-0166099.
PR   17-NOV-1999;   99US-0166369.
PR   23-DEC-1999;   99US-0171900.
PR   23-DEC-1999;   99US-0171901.
PR   23-DEC-1999;   99US-0171902.
PR   11-FEB-2000; 2000US-0181749.
PR   14-MAR-2000; 2000US-0189258.
PR   14-MAR-2000; 2000US-0189259.
PR   10-APR-2000; 2000US-0195898.
PR   10-APR-2000; 2000US-0195899.
PR   10-APR-2000; 2000US-0196078.
PR   28-APR-2000; 2000US-0200419.
PR   12-MAY-2000; 2000US-0203630.
PR   12-JUN-2000; 2000US-0210741.
PR   12-JUN-2000; 2000US-0210982.
PR   21-AUG-2000; 2000US-0226760.
PR   26-SEP-2000; 2000US-0235418.
PR   26-SEP-2000; 2000US-0235779.
PR   20-OCT-2000; 2000US-0242332.
PR   20-OCT-2000; 2000US-0242343.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Chen R,  Dang HT,  Lowitz KP;
XX
DR   WPI; 2001-355616/37.
DR   N-PSDB; AAS08271.
XX
PT   Endogenous and non-endogenous versions of human G-protein coupled
PT   receptors for direct identification of candidate compounds as agonists,
PT   inverse agonists or partial agonists for use as therapeutic agents -
XX
PS   Example 5; Page 146-149; 160pp; English.
XX
CC   The sequence is a G-protein coupled receptor (GPCR) fusion protein,
CC   hRUP13-Gs, being the human hRUP13 fused to the rat Gs protein.
CC   The endogenous and non-endogenous, constitutively activated versions
CC   of human G-protein coupled receptors (GPCR), are useful for direct
CC   identification of candidate compounds as receptor agonists, inverse
CC   agonists or partial agonists having applicability as therapeutic
CC   agents for treating diseases related to GPCR, e.g. lung cancer.
CC   Non-endogenous version of human GPCRs are also utilized in research
CC   settings and in vitro and in vivo system, incorporating GPCRs can be
CC   utilised to elucidate and understand the roles these receptors
CC   play in the human condition, both normal and diseased.
XX
SQ   Sequence   869 AA;

  Query Match             100.0%;  Score 2318;  DB 22;  Length 869;
  Best Local Similarity   100.0%;  Pred. No. 6e-228;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy     61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy    121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
          |||||||||||||||||||||||||||||||
Db    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
```

RESULT 15
AAU11896
ID AAU11896 standard; Protein; 451 AA.
XX
AC AAU11896;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human novel G protein-coupled receptor, GPCR8b.
XX
KW Human; GPCR8b; G protein coupled receptor; developmental disease;
KW immune disease; taste and scent detectability disorder;
KW Burkitt's lymphoma; signal transduction pathway disorder;
KW retinal disease; cell growth rate disorder; feeding disorder;
KW control of feeding; obesity; starvation; fungal infection;
KW noninsulin-dependent diabetes mellitus; bacterial infection;
KW protozoal infection; viral infection; pain; cancer; anorexia; bulimia;
KW asthma; Parkinson's disease; acute heart failure; hypotension;
KW hypertension; urinary retention; osteoporosis; Crohn's disease;
KW multiple sclerosis; Albright Hereditary Osteodystrophy; angina pectoris;
KW myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;
KW manic depression; delirium; dementia; severe metal retardation;
KW autosomal dominant (2) Acrocallosal syndrome; dyskinesia;
KW Huntington's disease; Gilles de la Tourette syndrome;
KW haematopoietic disorder; cell signal processing disorder;
KW metabolic pathway modulation disorder.
XX
OS Homo sapiens.
XX
PN WO200190187-A2.
XX
PD 29-NOV-2001.
XX
PF 24-MAY-2001; 2001WO-US17114.
XX
PR 24-MAY-2000; 2000US-206757P.
PR 25-MAY-2000; 2000US-207020P.
PR 28-JUN-2000; 2000US-214372P.
PR 19-JUL-2000; 2000US-219786P.
PR 25-JUL-2000; 2000US-220593P.
PR 10-OCT-2000; 2000US-239542P.
PR 18-DEC-2000; 2000US-256402P.
PR 26-FEB-2001; 2001US-271645P.
PR 09-MAR-2001; 2001US-274809P.
PR 13-MAR-2001; 2001US-275590P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Spytek KA, Majumder K, Tchernev VT, Grosse WM;
PI Szerkeres ED, Alsobrook JP, Burgess CE, Shimkets RA, Taupier RJ;
PI Casman SJ, Gangolli E, MacDougall JR, Stone DJ, Smithson G;
XX
DR WPI; 2002-106188/14.
DR N-PSDB; AAS19413.
XX
PT New polypeptide, useful for treating pain, cancer, urinary retention,
PT osteoporosis, Crohn's disease, dyskinesias, acute heart failure,
PT dementia, asthma and ulcer, comprises the isolated G-coupled
PT protein-receptor GPCRX -
XX
PS Claim 1; Page 37; 154pp; English.
XX
CC The invention relates to an isolated G protein-coupled receptor related
CC polypeptide (GPCR1a, 1b, 2, 3a, 3b, 4, 5, 6a, 6b, 7, 8a, 8b, and GPCR9)
CC and the polynucleotides encoding them. The GPCR, its encoding
CC polynucleotide and an anti-GPCR antibody are useful for diagnosing
CC treating or preventing a GPCR-associated disorder in a human e.g.
CC developmental diseases, immune diseases, taste and scent detectability
CC disorder, Burkitt's lymphoma, signal transduction pathway disorders,
CC retinal diseases including those involving photoreception, cell growth
CC rate disorders, feeding disorders, control of feeding, potential
CC obesity due to over-eating, potential disorders due to starvation,
CC noninsulin-dependent diabetes mellitus, bacterial, fungal, protozoal
CC and viral infections, pain, cancer, anorexia, bulimia, asthma,
CC Parkinson's disease, acute heart failure, hypotension, hypertension,
CC urinary retention, osteoporosis, Crohn's disease, multiple sclerosis,
CC Albright Hereditary Osteodystrophy, angina pectoris, myocardial
CC infarction, ulcers, allergies, benign prostatic hypertrophy, manic
CC depression, delirium, dementia, severe metal retardation, autosomal
CC dominant (2) Acrocallosal syndrome and dyskinesias, such as
CC Huntington's disease or Gilles de la Tourette syndrome, haematopoietic
CC disorders, and disorders related to cell signal processing and
CC metabolic pathway modulation. The present sequence represents GPCR8b.
XX
SQ Sequence 451 AA;

```
  Query Match             100.0%;  Score 2317;  DB 23;  Length 451;
  Best Local Similarity   99.8%;  Pred. No. 3.1e-228;
  Matches  450;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy     61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy    121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
```

```
Qy     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRVPGQIAE 420

Qy     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
           |||||||||||||||||||||||||||||||
Db     421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
```

Search completed: December 10, 2002, 09:57:21
Job time : 43 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: December 10, 2002, 09:55:55 ; Search time 35 Seconds
(without alignments)
2655.064 Million cell updates/sec

Title: US-09-838-028-2
Perfect score: 2318
Sequence: 1 MESSPIPQSSGNSSTLGRVP.....SDIIMSDSYLRPAASPRLES 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2318	100.0	451	4	Q8TDV4	Q8tdv4 homo sapien
2	435.5	18.8	368	4	Q8TAM0	Q8tam0 homo sapien
3	347	15.0	358	11	Q9QX37	Q9qx37 mus musculu
4	347	15.0	397	11	Q9D282	Q9d282 mus musculu
5	333.5	14.4	571	6	Q9TTM9	Q9ttm9 sus scrofa
6	330.5	14.3	405	6	Q9GJS6	Q9gjs6 ovis aries
7	329	14.2	407	6	Q9MZ00	Q9mz00 sus scrofa
8	328	14.2	438	11	Q63004	Q63004 rattus norv
9	327.5	14.1	405	6	Q9GJT0	Q9gjt0 ovis aries
10	327.5	14.1	405	6	Q9GL56	Q9gl56 ovis aries
11	326.5	14.1	405	6	Q9GL57	Q9gl57 ovis aries
12	322	13.9	399	5	Q9NG02	Q9ng02 apis mellif
13	318.5	13.7	402	13	Q90X46	Q90x46 brachydanio
14	316.5	13.7	559	11	Q9QW71	Q9qw71 rattus sp.
15	306	13.2	466	4	Q96RE8	Q96re8 homo sapien
16	304	13.1	515	11	Q9DBL0	Q9dbl0 mus musculu
17	302.5	13.1	464	5	Q9GQ54	Q9gq54 aedes aegyp
18	301	13.0	518	6	Q9MYI8	Q9myi8 oryctolagus
19	300	12.9	429	4	Q13729	Q13729 homo sapien
20	300	12.9	455	4	O60451	O60451 homo sapien
21	300	12.9	499	4	Q13675	Q13675 homo sapien
22	299.5	12.9	378	4	Q96KH9	Q96kh9 homo sapien
23	299	12.9	429	6	Q9MZU3	Q9mzu3 oryctolagus
24	297	12.8	466	4	Q9UD63	Q9ud63 homo sapien
25	296	12.8	425	6	Q9MZU2	Q9mzu2 oryctolagus
26	292.5	12.6	603	5	Q9VG57	Q9vg57 drosophila
27	291.5	12.6	387	4	Q96KI0	Q96ki0 homo sapien
28	290.5	12.5	447	6	Q8SPH2	Q8sph2 sus scrofa
29	289	12.5	446	6	Q8WND7	Q8wnd7 bos taurus
30	288.5	12.4	419	5	O77254	O77254 boophilus m
31	287	12.4	466	6	Q9TSW7	Q9tsw7 sus scrofa
32	284.5	12.3	385	5	Q24038	Q24038 drosophila
33	283	12.2	407	4	O75963	O75963 homo sapien
34	282	12.2	445	13	Q98841	Q98841 anguilla an
35	281.5	12.1	407	5	Q9VG54	Q9vg54 drosophila
36	280.5	12.1	437	13	O42316	O42316 cyprinus ca
37	279.5	12.1	446	13	O42315	O42315 cyprinus ca
38	279	12.0	402	5	O44198	O44198 apis mellif
39	278.5	12.0	422	6	Q9N296	Q9n296 pongo pygma
40	278.5	12.0	508	5	Q9VCZ3	Q9vcz3 drosophila
41	277.5	12.0	349	4	Q9UD67	Q9ud67 homo sapien
42	277.5	12.0	422	6	Q9N297	Q9n297 gorilla gor
43	277.5	12.0	474	11	Q8R456	Q8r456 rattus norv
44	276.5	11.9	422	6	Q9N298	Q9n298 pan troglod
45	276.5	11.9	445	13	Q98842	Q98842 anguilla an

ALIGNMENTS

RESULT 1
Q8TDV4
ID Q8TDV4 PRELIMINARY; PRT; 451 AA.
AC Q8TDV4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative G-protein coupled receptor.
GN GPCR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RT "Identification of G protein-coupled receptor genes from the human
RT genome sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB083585; BAB89298.1; -.
KW Receptor.
SQ SEQUENCE 451 AA; 49292 MW; 1F477F112E1CEA1A CRC64;

Query Match 100.0%; Score 2318; DB 4; Length 451;
Best Local Similarity 100.0%; Pred. No. 3.8e-194;
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
```

```
Qy    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240

Qy    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAE 420

Qy    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
          |||||||||||||||||||||||||||||||
Db    421 ETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
```

```
RESULT 2
Q8TAM0
ID   Q8TAM0      PRELIMINARY;      PRT;   368 AA.
AC   Q8TAM0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   G protein-coupled receptor 62.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC026357; AAH26357.1; -.
KW   Receptor.
SQ   SEQUENCE   368 AA;  37618 MW;  9D5D6A9D98D12FBA CRC64;

  Query Match             18.8%;  Score 435.5;  DB 4;  Length 368;
  Best Local Similarity   31.8%;  Pred. No. 6.9e-30;
  Matches  127;  Conservative   56;  Mismatches  144;  Indels   73;  Gaps   14;

Qy     40 ASE---SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPL 96
          |||   |: |    :::: |: || |:: |: :|| ||  ::: |||:|||||| ::|||
Db      9 ASEVAGSLGLILAAVVEVGALLGNGALLVVVLRTPGLRDALYLAHLCVVDLLAAASIMPL 68

Qy     97 AMLSSSAL-FDHALFGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMT 155
           :|::          |   ||   |||    :    | |:|: : ||  :|||:|   |
Db     69 GLLAAPPPGLGRVRLGPAPCRAARFLSAALLPACTLGVAALGLARYRLIVHPLRPGSRPP 128

Qy    156 LGLVASVLVGVWVKALAMASVPVLGRVSWEEGAPSVPPGCSLQWSHSAYCQL-------F 208
            |   ||  ||  |  : :: :|       | |  ||         | | :       |
Db    129 PVL---VLTAVWAAAGLLGALSLL-------GPPPAPPPA------PARCSVLAGGLGPF 172

Qy    209 VVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTS 268
            ::|:| | || ||:|  |  :| ||| ||::  | |      | ||:|| || :::
Db    173 RPLWALLAFALPALLLLGAYGGIFVVARRAALR-PPRPA--RGSRLRSDSLDSRLSIL-- 227

Qy    269 SGAPQTTPHRT-FGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVT 327
              | |:   ||||:      ||| |||||       | |       : |: ||
Db    228 ------PPLRSRLPGGKAALAPALAVGQFAACWLPY-----GCACLAPAARAAEAEAAVT 276

Qy    328 WIGYFCFTSNPFFYGCLNRQIR---GELSKQFVCFFKPAPEEELRLPSREGSIEENFLQF 384
          |: |  | ::|| || | | :|    | ||:: :    | |
Db    277 WVAYSAFAAHPFLYGLLQRPVRLALGRLSRRAL----PGPVR----------------- 314

Qy    385 LQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAEETSE 424
              |  ::|  | |    | ||      | :  |:| |
Db    315 ----ACTPQAWHPRALLQCLQRPPEGPAVGPSEAPEQTPE 350
```

```
RESULT 3
Q9QX37
ID   Q9QX37      PRELIMINARY;      PRT;   358 AA.
AC   Q9QX37;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Histamine H2 receptor.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129SVJ;
RA   Fukushima Y., Asano T., Sugano K.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF019138; AAD01634.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   358 AA;  40284 MW;  D270C2C66AB7A256 CRC64;

  Query Match             15.0%;  Score 347;  DB 11;  Length 358;
  Best Local Similarity   27.6%;  Pred. No. 3.6e-22;
  Matches  108;  Conservative   58;  Mismatches  152;  Indels   74;  Gaps   12;

Qy     42 ESVALFFMLLLDLT-----AVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLM 94
          :|:||   : : ||      ||||  |   ::    ||     |: |   |||  | :|
Db     13 DSIALKVTISVVLTTLIFITVAGNVVVCLAVSLNRRLRSLTNCFIVSLAATDLLLGLLVM 72

Qy     95 PLAMLSSSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYE 151
          |      ||::  :    ||:| | :|  | |   : :||::  |:::||  |  |:||
Db     73 PF-----SAIYQLSFKWSFGQVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYP 127

Qy    152 VRMTLGLVASVLVGVWVKALAMASVPVLGRVSWEE------GAPSVPPGCSLQWSHSAYC 205
          | :|  ||  || :|| :: :: : :   : |       |  :    | :| :
Db    128 VLVTPVRVAISLVFIWVISITLSFLSI--HLGWNSRNGTRGGNDTFK--CKVQVN----- 178

Qy    206 QLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTM 265
          ::: :|  :: | ||||:: | |  :|:::||  | :   : :|
Db    179 EVYGLVDGMVTFYLPLLIMCVTYYRIFKIAREQAKRINHISSWKAA------------- 224

Qy    266 VTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESV 325
                    |    || | | || | |::|| |||:  :|  |         || :
Db    225 ------------TIREHKATVTLAAVMGAFIVCWFPYFTAFVYRGLRGDDAVNEVVEGI 271

Qy    326 VTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFFKP--APEEELRL--------PSREG 375
          | |:||    ||  |  |||  |    : | |      : :  |||         ||||
Db    272 VLWLGYANSALNPILYATLNRDFRMAYQQLFHCKLASHNSHKTSLRLNNSLLSRSQSREG 331

Qy    376 SIEENFLQFLQGTGCPSESWVSRPLPSPKQEP 407
            :|     ||         |    |  |:  |
Db    332 RWQEEKPLKLQ-------VWSGTELTHPQGSP 356
```

```
RESULT 4
Q9D282
ID   Q9D282      PRELIMINARY;      PRT;   397 AA.
AC   Q9D282;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Histamine receptor H 2.
GN   HRH2.
```

```
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=CECUM;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AK020259; BAB32044.1; -.
DR   MGD; MGI:108482; Hrh2.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01570; NPFFRECEPTOR.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ   SEQUENCE   397 AA;  44793 MW;  BB4719084F4BDD59 CRC64;

  Query Match             15.0%;  Score 347;  DB 11;  Length 397;
  Best Local Similarity   27.6%;  Pred. No. 4e-22;
  Matches  108;  Conservative   58;  Mismatches  152;  Indels   74;  Gaps   12;

Qy      42 ESVALFFMLLLDLT-----AVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLM 94
           :|:||   : : ||       ||||  |   ::    |'     |:  |   |||  | :|
Db      13 DSIALKVTISVVLTTLIFITVAGNVVVCLAVSLNRRLRSLTNCFIVSLAATDLLLGLLVM 72

Qy      95 PLAMLSSSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYE 151
           |      ||::  :    ||:| | :|  | |   : :||::  |:::||  |  |:||
Db      73 PF-----SAIYQLSFKWSFGQVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYP 127

Qy     152 VRMTLGLVASVLVGVWVKALAMASVPVLGRVSWEE------GAPSVPPGCSLQWSHSAYC 205
           | :|   ||  || :|| :: :: : :   : |       |  :    | :| :
Db     128 VLVTPVRVAISLVFIWVISITLSFLSI--HLGWNSRNGTRGGNDTFK--CKVQVN----- 178

Qy     206 QLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTM 265
           ::: :|  :: | ||||:: | |  :|::||  | :   : :|
Db     179 EVYGLVDGMVTFYLPLLIMCVTYYRIFKIAREQAKRINHISSWKAA-------------- 224

Qy     266 VTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESV 325
                  |    || | | || | |::|| |||:  :|  |          || :
Db     225 ------------TIREHKATVTLAAVMGAFIVCWFPYFTAFVYRGLRGDDAVNEVVEGI 271

Qy     326 VTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFFKP--APEEELRL--------PSREG 375
           | |:||     ||  |  |||  |    : | |      : :  |||         ||||
Db     272 VLWLGYANSALNPILYATLNRDFRMAYQQLFHCKLASHNSHKTSLRLNNSLLSRSQSREG 331

Qy     376 SIEENFLQFLQGTGCPSESWVSRPLPSPKQEP 407
             :|     ||        |    |  |:  |
Db     332 RWQEEKPLKLQ-------VWSGTELTHPQGSP 356
```

```
RESULT 5
Q9TTM9
ID   Q9TTM9      PRELIMINARY;      PRT;   571 AA.
AC   Q9TTM9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Alpha-1D adrenergic receptor.
GN   ALPHA-1D.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Uhlen S., Wraith A.;
RT   "Characterization of the pig alpha-1D adrenergic receptor.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AJ250492; CAB59347.1; -.
DR   EMBL; AJ250493; CAB59347.1; JOINED.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   571 AA;  60698 MW;  BACFFB5C903972B1 CRC64;

  Query Match             14.4%;  Score 333.5;  DB 6;  Length 571;
  Best Local Similarity   24.0%;  Pred. No. 9e-21;
  Matches  125;  Conservative   85;  Mismatches  201;  Indels  109;  Gaps   17;

Qy       6 IPQSSGNSSTLGRVPQ--------TPGPSTA----SGVPEVGLRDVASESVAL-FFMLLL 52
           :| ::|:   :|             || : |    :|   ||   |::: | :  |:
Db      46 VPGTAGSGGVVGAGSDENNRSSAGEPGAAGAGGEVNGTAAVGGLVVSAQGVGVGVFLAAF 105

Qy      53 DLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFDHALF 110
            | |||||  |:  :|    |:     |: :| : ||| : |::| :  ::  :     |
Db     106 ILMAVAGNLLVILSVACNRHLQTVTNYFIVNLAVADLLLSATVLPFS--ATMEVLGFWAF 163

Qy     111 GEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKA 170
           |   | ::  : |   : :|||:  |:|:||  | | ::|   ||    |::|  :|  |
Db     164 GRAFCDVWAAVDVLCCTASILSLCTISVDRYVGVRHSLKYPSIMTERKAAAILALLWAVA 223

Qy     171 LAMASVPVLGRVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLIL 225
           : ::  |:||   |:|    |||          :| :     : |  ::  | ||: :|:
Db     224 IVVSVGPLLG---WKE---PVPP-------DERFCGITEEAGYAVFSSLCSFYLPMAVIV 270

Qy     226 VVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSS-------GAPQTTPHR 278
           |:|| :: |||         |   ::  | ::  :  |     ||        || ::|
Db     271 VMYCRVYVVARSTTRS---LEAGVKRERGKASEVVLRIHCRGSSTGTDRGHGAMRSTKGH 327

Qy     279 TFGGG------------KAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVV 326
           ||               |||  |  | | |:||| |:|       :|  |   :  |  |:
Db     328 TFRSSLSLRLLKFSREKKAAKTLAIVVGVFVLCWFPFFFVLPLGSLFPQLKPSEGVFKVI 387

Qy     327 TWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFFK----------------------PA 363
           |:|||    ||  | | :|: :    :   |                         |
Db     388 FWLGYFNSCVNPLIYPCSSREFKRAFLRLLRCQCHHSRRRRRPLWRAYGHHWLASNGGPR 447

Qy     364 P-----------EEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAV-D 411
           |           |  | ||: |             |  ||      |:   :: | :  |
Db     448 PDCAPGLGAAPREAPLALPAPE-----------ATDTPSAPEAQAPVVGRRKPPYSFRD 495

Qy     412 FRIPGQIAEETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
           :|: |    |:     || : :      :|  ::|  |:
Db     496 WRLLGPFRRPTT-----QLRAKVSSLSQKIRAGSAPCAEA 530

RESULT 6
```

```
Q9GJS6
ID   Q9GJS6      PRELIMINARY;      PRT;   405 AA.
AC   Q9GJS6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Beta 3 adrenergic receptor.
GN   B3AR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Forrest R.H., Hickford J.G.H.;
RT   "Polymorphism within the ovine beta 3 adrenergic receptor gene.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF314204; AAG31167.1; -.
DR   EMBL; AF314202; AAG31165.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   405 AA;  42928 MW;  A0DD3AEBAF814E75 CRC64;

  Query Match             14.3%;  Score 330.5;  DB 6;  Length 405;
  Best Local Similarity   29.4%;  Pred. No. 1.1e-20;
  Matches  116;  Conservative   62;  Mismatches  170;  Indels   47;  Gaps   17;

Qy       7 PQSSGNSSTLGRVPQTP--GPST--ASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAA 62
           |   |||| |   |  |   |:|  |||:| |      : ::|   : |  |  | ||
Db       3 PWPPGNSS-LTPWPDIPTLAPNTANASGLPGVPW----AVALAGALLALAVLATVGGNLL 57

Qy      63 VMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           |:  ||:|| |:    |||  |   ||:  | ::|    :: ||  |   |   | |:
Db      58 VIVAIARTPRLQTMTNVFVTSLATADLVVGLLVVPPG--ATLALTGHWPLGVTGCELWTS 115

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           : |  |: :| :: |: |:||  | :|:||   :|     : :| ||| : |::  |::
Db     116 VDVLCVTASIETLCALAVDRYLAVTNPLRYGALVTKRRARAAVVLVWVVSAAVSFAPIMS 175

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA 235
           :  |  || :    |    |:   |       : :: : : | ||||::| ||  :| ||
Db     176 K-WWRVGADAEAQRCH---SNPRCCTFASNMPYALLSSSVSFYLPLLVMLFVYARVFVVA 231

Qy     236 ----RVAAMQHGPLPTWMETPR-QRSESLSSRSTMVTSSGAP-------QTTPHRTFGGG 283
               |:   : |  |     |   || |        : :| |       :  | |
Db     232 TRQLRLLRRELGRFPPEESPPAPSRSGSPGPAGPYASPAGVPSYGRRPARLLPLREH--- 288

Qy     284 KAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGC 343
           :|   |  : | | |||||:|  ::  ||    : :|     : |:||     ||  | |
Db     289 RALRTLGLIMGTFTLCWLPFFVVNVVRALGGPSLVSGPTFLALNWLGYANSAFNPLIY-C 347

Qy     344 LNRQIRGELSKQFVCFFKPAPEEEL------RLPS 372
           :   |    :: :|  :  ||| |      | ||
Db     348 RSPDFRSAF-RRLLC--RCPPEEHLAAASPPRAPS 379


RESULT 7
Q9MZ00
ID   Q9MZ00      PRELIMINARY;      PRT;   407 AA.
AC   Q9MZ00;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Beta-3-adrenergic receptor.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smith T.R., Bidwell C.A., Mills S.E.;
RT   "Sus scrofa beta-3-adrenergic receptor (BAR3) gene.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF274007; AAF82301.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   407 AA;  43610 MW;  C6598382A9B38DD9 CRC64;

  Query Match             14.2%;  Score 329;  DB 6;  Length 407;
  Best Local Similarity   28.6%;  Pred. No. 1.5e-20;
  Matches  125;  Conservative   61;  Mismatches  177;  Indels   74;  Gaps   21;

Qy       4 SPIPQSSGNSSTLGRVPQTPGPST-------ASGVPEVGLRDVASESVALFFMLLLD--L 54
             :| ||  ||||    :|  |  ||         ||:| |         :|||   ||    |
Db       2 APWPQ--GNSS----LPPRPDVSTLAPNNANTSGLPGV------PWAVALAGALLAPAVL 49

Qy      55 TAVAGNAAVMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGE 112
             | ||  |:  ||:|| |:    |||  |   ||:  | ::|    :: ||  |   |
Db      50 ATVGGNLLVIVAIARTPRLQTMTNVFVTSLATADLVVGLLVVPPG--TTLALTGHWPLGA 107

Qy     113 VACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALA 172
           | |:  : |  |: :| :: |: |:||  | :|:||   :|     : :| ||| : |
Db     108 TCCELWTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALVTKRRARAAVVLVWVVSAA 167

Qy     173 MASVPVLGRVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVV 227
           ::  |:: :  |  || :    |    |: : |       : :: : : | ||||::| |
Db     168 VSFAPIMSK-WWRVGADAEAQRCH---SNPSCCTFASNMPYALLSSSVSFYLPLLVMLFV 223

Qy     228 YCSMFRVARVAAMQHGPLPTW--METPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKA 285
           |  :|    | |     |  |     | : |    |||    : | |  :|      |:
Db     224 YARVF----VVATSQLRLLRWELSRFPPEESPPAPSRS-QSPAPGRPWPSPAGVPSHGRR 278

Qy     286 AVVLL-----------AVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFC 333
             ||           : | | |||||:|  ::  ||    :          : |:||
Db     279 PARLLPLREHRALCTLGLIMGTFTLCWLPFFVVNVVRALGGPSLVPVPAFLALNWLGYAN 338

Qy     334 FTSNPFFYGCLNRQIRGELSKQFVCFFKPAPEEELRL--PSREGSIEENFLQFLQGTGCP 391
             ||  | | :    |    :: :|  :  ||| |    | |  |            | |
Db     339 SAFNPLIY-CHSPDFRSAF-RRLLC--RCGPEEHLAAASPPRAPS-----------GAP 382

Qy     392 SESWVSRPLPSPKQEPP 408
             :: |  | :| ||
Db     383 ET--LTHPAES-RQSPP 396


RESULT 8
Q63004
ID   Q63004      PRELIMINARY;      PRT;   438 AA.
AC   Q63004;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   5-HT6 serotonin receptor.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=STRIATUM;
RX   MEDLINE=96102917; PubMed=8522988;
```

```
RA   Kohen R., Metcalf M.A., Khan N., Druck T., Huebner K., Lachowicz J.E.,
RA   Meltzer H.Y., Sibley D.R., Roth B.L., Hamblin M.W.;
RT   "Cloning, characterization, and chromosomal localization of a human 5-
RT   HT6 serotonin receptor.";
RL   J. Neurochem. 66:47-56(1996).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; L41146; AAA92633.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   438 AA;  46786 MW;  DB90403892F691A9 CRC64;

  Query Match           14.2%;  Score 328;  DB 11;  Length 438;
  Best Local Similarity  27.4%;  Pred. No. 2e-20;
  Matches  129;  Conservative   61;  Mismatches  184;  Indels   96;  Gaps   19;

Qy       2 ESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNA 61
Db       4 EPGPV-----NSST---PAWGPGPPPAPG---------GSGWVAAALCVVIVLTA-AANS 45

Qy      62 AVMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYL 119
Db      46 LLIVLICTQPAVRNTSNFFLVSLFTSDLMVGLVVMPPAML--NALYGRWVLARGLCLLWT 103

Qy     120 FLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVL 179
Db     104 AFDVMCCSASILNLCLISLDRYLLILSPLRYKLRMTAPRALALILGAWSLAALASFLPLL 163

Qy     180 GRVSWEE-GAPSVP-PGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMF 232
Db     164 --LGWHELGKARTPAPG---------QCRLLASLPFVLVASGVTFFLPSGAICFTYCRIL 212

Qy     233 RVARVAAMQHGPLPTWM----------ETPRQRSESLSSRSTMVTSSGAPQTTPHRTFG 281
Db     213 LAARKQAVQVASLTTGTAGQALETLQVPRTPRPGMESADSRRLATKHSRKAL-------- 264

Qy     282 GGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFY 341
Db     265 --KASLTLGILLGMFFVTWLPFFVANIAQAV-CDCISPGLFD-VLTWLGYCNSTMNPIIY 320

Qy     342 GCLNRQIRGELSKQFVCFFKPAPEEE-------------------------LRLP---- 371
Db     321 PLFMRDFKRALGRFLPCVHCP-PEHRASPASPSMWTSHSGARPGLSLQQVLPLPLPPNSD 379

Qy     372 SREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAEE 421
Db     380 SDSASGGTSGLQLTAQLLLPGEA--TRDPPPPTRATTVVNFFVTDSVEPE 427


RESULT 9
Q9GJT0
ID   Q9GJT0      PRELIMINARY;      PRT;   405 AA.
AC   Q9GJT0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Beta 3 adrenergic receptor.
GN   B3AR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Forrest R.H., Hickford J.G.H.;
RT   "Ovis aries beta 3 adrenergic receptor (B3AR) gene - allele F,
RT   complete cds.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Forrest R.H., Hickford J.G.H.;
RT   "Polymorphism within the ovine beta 3 adrenergic receptor gene.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF314205; AAG31168.1; -.
DR   EMBL; AF314201; AAG31164.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   405 AA;  42956 MW;  6C9FB1EBAF93C578 CRC64;

  Query Match           14.1%;  Score 327.5;  DB 6;  Length 405;
  Best Local Similarity  29.4%;  Pred. No. 2.1e-20;
  Matches  116;  Conservative   61;  Mismatches  171;  Indels   47;  Gaps   17;

Qy       7 PQSSGNSSTLGRVPQTP--GPST--ASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAA 62
Db       3 PWPPGNSS-LTPWPDIPTLAPNTANASGLPGVPW----AVALAGALLALAVLATVGGNLL 57

Qy      63 VMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
Db      58 VIVAIARTPRLQTMTNVFVTSLATADLVVGLLVVPPG--ATLALTGHWPLGVTGCELWTS 115

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
Db     116 VDVLCVTASIETLCALAVDRYLAVTNPLRYGALVTKRRARAAVVLVWVVSAAVSFAPIMS 175

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA 235
Db     176 K-WWRVGADAEAQRCH---SNPRCCTFASNMPYALLSSSVSFYLPLLVMLFVYARVFVVA 231

Qy     236 ----RVAAMQHGPLPTWMETPR-QRSESLSSRSTMVTSSGAP-------QTTPHRTFGGG 283
Db     232 TRQLRLLRRELGRFPPEESPPAPSRSGSPGPAGPYASPVGVPSYGRRPARLLPLREH--- 288

Qy     284 KAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGC 343
Db     289 RALRTLGLIMGTFTLCWLPFFVVNVVRALGGPSLVSGPTFLALNWLGYANSAFNPLIY-C 347

Qy     344 LNRQIRGELSKQFVCFFKPAPEEEL------RLPS 372
Db     348 RSPDFRSAF-RRLLC--RCPPEEHLAAASPPRAPS 379


RESULT 10
Q9GL56
ID   Q9GL56      PRELIMINARY;      PRT;   405 AA.
AC   Q9GL56;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Beta 3 adrenergic receptor.
GN   B3AR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Forrest R.H., Hickford J.G.H.;
RT   "Polymorphism within the ovine beta 3 adrenergic receptor gene.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
```

```
DR   EMBL; AF314203; AAG31166.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   405 AA;  42886 MW;  EF4B13BB413EA413 CRC64;

  Query Match             14.1%;  Score 327.5;  DB 6;  Length 405;
  Best Local Similarity   29.1%;  Pred. No. 2.1e-20;
  Matches  115;  Conservative   62;  Mismatches  171;  Indels   47;  Gaps   17;

Qy       7 PQSSGNSSTLGRVPQTP--GPST--ASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAA 62
           |   |||| |   |  |   |:|  |||:| |      : ::|   : |  |    ||
Db       3 PWPPGNSS-LTPWPDIPTLAPNTANASGLPGVPW----AVALAGALLALAVLATAGGNLL 57

Qy      63 VMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           |:  ||:|| |:    |||  |   ||:  | ::|    :: ||  |   |   | |:
Db      58 VIVAIARTPRLQTMTNVFVTSLATADLVVGLLVVPPG--ATLALTGHWPLGVTGCELWTS 115

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           : |  |: :| :: |: |:||  | :|:||   :|     : :| ||| : |::  |::
Db     116 VDVLCVTASIETLCALAVDRYLAVTNPLRYGALVTKRRARAAVVLVWVVSAAVSFAPIMS 175

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA 235
           :  |  || :   |    |:  |       : :: : : | ||||::| ||  :| ||
Db     176 K-WWRVGADAEAQRCH---SNPRCCTFASNMPYALLSSSVSFYLPLLVMLFVYARVFVVA 231

Qy     236 ----RVAAMQHGPLPTWMETPR-QRSESLSSRSTMVTSSGAP-------QTTPHRTFGGG 283
               |:   : |  |      |   || |        : :| |       :  | |
Db     232 TRQLRLLRRELGRFPPEESPPAPSRSGSPGPAGPYASPAGVPSYGRRPARLLPLREH--- 288

Qy     284 KAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGC 343
           :|   |  : | | ||||||:|  ::  ||    : :|     : |:||     ||  | |
Db     289 RALRTLGLIMGTFTLCWLPFFVVNVVRALGGPSVVSGPTFLALNWLGYANSAFNPLIY-C 347

Qy     344 LNRQIRGELSKQFVCFFKPAPEEEL------RLPS 372
             :   |    :: :|  :  ||| |      | ||
Db     348 RSPDFRSAF-RRLLC--RCPPEEHLAAASPPRAPS 379


RESULT 11
Q9GL57
ID   Q9GL57      PRELIMINARY;      PRT;   405 AA.
AC   Q9GL57;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Beta 3 adrenergic receptor.
GN   B3AR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Forrest R.H., Hickford J.G.H.;
RT   "Polymorphism within the ovine beta 3 adrenergic receptor gene.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF314200; AAG31163.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   405 AA;  42928 MW;  6C9FB1EBAF90F54B CRC64;

  Query Match             14.1%;  Score 326.5;  DB 6;  Length 405;
  Best Local Similarity   29.3%;  Pred. No. 2.5e-20;
  Matches  114;  Conservative   61;  Mismatches  173;  Indels   41;  Gaps   16;

Qy       7 PQSSGNSSTLGRVPQTP--GPST--ASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAA 62
           |   |||| |   |  |   |:|  |||:| |      : ::|   : |  |    ||
Db       3 PWPPGNSS-LTPWPDIPTLAPNTANASGLPGVPW----AVALAGALLALAVLATVGGNLL 57

Qy      63 VMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           |:  ||:|| |:    |||  |   ||:  | ::|    :: ||  |   |   | |:
Db      58 VIVAIARTPRLQTMTNVFVTSLATADLVVGLLVVPPG--ATLALTGHWPLGVTGCELWTS 115

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           : |  |: :| :: |: |:||  | :|:||   :|     : :| ||| : |::  |::
Db     116 VDVLCVTASIETLCALAVDRYLAVTNPLRYGALVTKRRARAAVVLVWVVSAAVSFAPIMS 175

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA 235
           :  |  || :   |    |:  |       : :: : : | ||||::| ||  :| ||
Db     176 K-WWRVGADAEAQRCH---SNPRCCTFASNMPYALLSSSVSFYLPLLVMLFVYARVFVVA 231

Qy     236 ----RVAAMQHGPLPTWMETPR-QRSESLSSRSTMVTSSGAP-------QTTPHRTFGGG 283
               |:   : |  |      |   || |        : :| |       :  | |
Db     232 TRQLRLLRRELGRFPPEESPPAPSRSGSPGPAGPYASPVGVPSYGRRPARLLPLREH--- 288

Qy     284 KAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGC 343
           :|   |  : | | ||||||:|  ::  ||    : :|     : |:||     ||  | |
Db     289 RALRTLGLIMGTFTLCWLPFFVVNVVRALGGPSLVSGPTFLALNWLGYANSAFNPLIY-C 347

Qy     344 LNRQIRGELSKQFVCFFKPAPEEELRLPS 372
             :   |    :: :|  :  ||| |   |
Db     348 RSPDFRSAF-RRLLC--RCPPEEHLAAAS 373


RESULT 12
Q9NG02
ID   Q9NG02      PRELIMINARY;      PRT;   399 AA.
AC   Q9NG02;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   G-protein coupled receptor.
GN   TYR1.
OS   Apis mellifera (Honeybee).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC   Apoidea; Apidae; Apis.
OX   NCBI_TaxID=7460;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20155587; PubMed=10693920;
RA   Blenau W., Balfanz S., Baumann A.;
RT   "Amtyr1: characterization of a gene from honeybee (Apis mellifera)
RT   brain encoding a functional tyramine receptor.";
RL   J. Neurochem. 74:900-908(2000).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AJ245824; CAB76374.1; -.
DR   HSSP; P29274; 1MMH.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ   SEQUENCE   399 AA;  44650 MW;  F5FE4E4F13336433 CRC64;

  Query Match             13.9%;  Score 322;  DB 5;  Length 399;
  Best Local Similarity   25.6%;  Pred. No. 6.1e-20;
  Matches  100;  Conservative   64;  Mismatches  139;  Indels   88;  Gaps   12;
```

```
Qy      23 PGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAAVMAVIAKTPALR--KFVFVF 80
           |   | | :|   : : |: |: | |::   |  | ||| |:  :     ||  :  |:
Db      23 PEEETGSNLP---VWEAAAASLTLGFLV---LATVLGNALVILSVFTYRPLRIVQNFFIV 76

Qy      81 HLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILSVSAINVER 140
           | : ||  |: :||  :  :  |    :||   |:|:|   |   : :||:: || ::|
Db      77 SLAVADLAVAILVMPFNV--AYLLLGKWIFGIHLCKLWLTCDVLCCTASILNLCAIALDR 134

Qy     141 YYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLGRVSW-EEGAPSVPPGCSLQW 199
           |: :  |: |  : ||  | : : |||: : |::| |: |   | ||  |  |
Db     135 YWAITDPINYAQKRTLKRVLATIAGVWILSGAISSPPLAGWNDWPEELEPGTP------- 187

Qy     200 SHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMF-----------RVARVAAMQ-- 241
             |||     :|:  ::  | :||||: :||  ::           | :|: |:|
Db     188 -----CQLTRRQGYVIYSSLGSFFIPLLLMSLVYLEIYLATRRRLRERARQSRINAVQST 242

Qy     242 ----------------HGPLPTWMETPRQRSES--LSSRSTMVTSSGAPQTTPHRTFGG 282
                          |       |||  ::   :    | ||  |   |:  || |
Db     243 RHREADDAEESVSSETNHNE----RSTPRSHAKPSLIDDEPTEVTIGGGGTTSSRRTTGS 298

Qy     283 ----------------------GKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPIS 318
                                 :||  |  : | |::|||||:|  :: |
Db     299 RAAATTTTVYQFIEERQRISLSKERRAARTLGVIMGVFVVCWLPFFLMYVIVPFCPDCCP 358

Qy     319 TGQVESVVTWIGYFCFTSNPFFYGCLNRQIR 349
           : ::   :||:||     ||  |   |   |
Db     359 SDRMVYFITWLGYVNSALNPLIYTIFNLDYR 389


RESULT 13
Q90X46
ID   Q90X46      PRELIMINARY;      PRT;   402 AA.
AC   Q90X46;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   SC:bZ20I5.4 (novel protein similar to human G-protein coupled receptor
DE   RE2).
GN   SC:BZ20I5.4.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lloyd D.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL590146; CAC94897.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   402 AA;  44373 MW;  5FB8875874267F00 CRC64;

  Query Match             13.7%;  Score 318.5;  DB 13;  Length 402;
  Best Local Similarity   27.7%;  Pred. No. 1.2e-19;
  Matches   96;  Conservative   69;  Mismatches  145;  Indels   37;  Gaps   11;

Qy      26 STASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAAVMAVIAKTPAL--RKFVFVFHLC 83
           :: :|: : ||  :  |||::  : :|   |  ||  ::  : | | |      ||| |
Db      12 NSTNGLDDNGL--MVLESVSIIIIAIL---ACLGNLVIVVTLYKKPYLLTPSNKFVFSLT 66

Qy      84 LVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILSVSAINVERYYY 143
             :|| :: ::|  : ||  :    :|| | |      | :   | ::|:: || ::|||
Db      67 SSNLLLSVLMLPFVVASS--VRRDWMFGVVWCNFTALLHLLVSSSSMLTLGAIAIDRYYA 124

Qy     144 VVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG-------RVSWEEGAPSVPPGCS 196
           |::|| | :::|       :| :|: :|     |: |       | |           |:
Db     125 VLYPMIYPMKITGNRAVLAIVYIWLHSLVGCLPPLFGWSSFEFDRFKWT---------CT 175

Qy     197 LQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAA--MQHGPLPTWMETPRQ 254
           : |        | | :     ||||: :|| |  :|||||: |  :  | :    |
Db     176 VSWHKEISYTAFWVTWCC---LLPLVAMLVCYGVIFRVARIKARKVYCGSVVVSQEESSS 232

Qy     255 RSESLSSRSTMVTSSGAPQTTPHRTFGGG--KAAVVLLAVGGQFLLCWLPYFSFHLYVAL 312
           ::    : :|  :|||: ::     : |   || : :| | | ||  | ||       ||
Db     233 QNNGRKNSNTSTSSSGSRKSL---IYSGSQCKAFITILVVLGTFLTTWGPYVVVISTEAL 289

Qy     313 SAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCF 359
           :   : |||::|:|: :     :|  ||  |: :| ||    :||
Db     290 LGKNSVSPQVETLVSWLSFTSAVCHPLIYGLWNKTVRKELLG--MCF 334


RESULT 14
Q9QW71
ID   Q9QW71      PRELIMINARY;      PRT;   559 AA.
AC   Q9QW71;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Alpha 1-adrenergic receptor subtype alpha 1D, alpha 1D-AR.
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92100054; PubMed=1661838;
RA   Perez D.M., Piascik M.T., Graham R.M.;
RT   "Solution-phase library screening for the identification of rare
RT   clones: isolation of an alpha 1D-adrenergic receptor cDNA.";
RL   Mol. Pharmacol. 40:876-883(1991).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 QF G-PROTEIN COUPLED RECEPTORS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ   SEQUENCE   559 AA;  59739 MW;  37CD9BD696D3A47D CRC64;

  Query Match             13.7%;  Score 316.5;  DB 11;  Length 559;
  Best Local Similarity   26.1%;  Pred. No. 2.7e-19;
  Matches  102;  Conservative   76;  Mismatches  156;  Indels   57;  Gaps   15;

Qy       6 IPQSSGNSSTL----GRVPQT----PGPSTA---SGVPEVGLRDVASESVAL-FFMLLLD 53
           :| ::|  : :    | : |:    || : :   :|   ||   |::: | :  |:
Db      42 VPGATGGGAVVGTGSGEINQSSTGEPGAAASGEVNGSAAVGGLVVSAQGVGVGVFLAAFI 101

Qy      54 LTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFDHALFG 111
           |||||||  |:  :|     |:     |: :| : ||| :  ::|  :  ::   :     ||
Db     102 LTAVAGNLLVILSVACNRHLQTVTNYFIVNLAVADLLLSAAVLPFS--ATMEVLGFWAFG 159

Qy     112 EVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKAL 171
             | ::  : |   : :|||:  |:|:||  | | ::|   |     |::|  :|  ||
Db     160 RTFCDVWAAVDVLCCTASILSLCTISVDRYVGVRHSLKYPAIMIERKAAAILALLWAVAL 219

Qy     172 AMASVPVLGRVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILV 226
           ::  |:||   |:|    |||          :| :     : :  :|  | ||: :|:|
Db     220 VVSVGPLLG---WKE---PVPP-------DERFCGITEEVGYAIFSSVCSFYLPMAVIVV 266

Qy     227 VYCSMFRVAR--VAAMQHGPLPTWMETPRQRSE---SLSSRSTMVTSSGAP--QTTPHRT 279
           :|| :: |||    ::: |       | : ||    :  |    :: | |  |::   |
Db     267 MYCRVYVVARSTTRSLEAG----IKREPGKASEVVLRIHCRGARTSAKGYPGTQSSKGHT 322

Qy     280 FGGG-----------KAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVT 327
                         |||  |  | | |:||| |:|      ::  |   :  |  |:
Db     323 LRSSLSVRLLKFSREKKAAKTLAIVVGVFVLCWFPFFFVLPLGSIFPQIKPSEGVFKVIF 382

Qy     328 WIGYFCFTSNPFFYGCLNRQIRGELSKQFVC 358
           |:|||    ||  | | :|: :    :   |
Db     383 WLGYFNSCVNPLIYPCSSREFKRAFLRLLRC 413
```

```
RESULT 15
Q96RE8
ID   Q96RE8      PRELIMINARY;      PRT;   466 AA.
AC   Q96RE8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Adrenergic receptor alpha-1a.
GN   ADRA1A.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Banerjee A.G.N., Aarti A.;
RT   "RT-PCR cloning and sequence analysis of adrenergic receptor subtype-
RT   alpha-1a cDNA from human prostrate cell-line DU-145.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF395806; AAK77197.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   466 AA;  51431 MW;  876CBFB3E323B7A1 CRC64;

  Query Match             13.2%;  Score 306;  DB 4;  Length 466;
  Best Local Similarity   27.8%;  Pred. No. 1.8e-18;
  Matches  110;  Conservative   62;  Mismatches  133;  Indels   90;  Gaps   18;

Qy      10 SGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLL--LDLTAVAGNAAVMAVI 67
           |||:|      | | |   |                  |  ::|  | |  | ||  |:  :
Db       5 SGNASDSSNCTQPPAPVNIS------------KAILLGVILGGLILFGVLGNILVILSV 51

Qy      68 AKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFD---HALFGEVACRLYLFLS 122
           |    |      :: :| : |||   |::|      ||:|:   :  || | | ::  :
Db      52 ACHRHLHSVTHYYIVNLAVADLLLTSTVLPF-----SAIFEVLGYWAFGRVFCNIWAAVD 106

Qy     123 VCFVSLAILSVSAINVERYYYVVHPMRYEVRMT--LGLVASVLVGVWVKALAMASVPVLG 180
           |   : :|: :  |:::||  | ||:||   :|   ||:|  |: ||  :| ::  |: |
Db     107 VLCCTASIMGLCIISIDRYIGVSHPLRYPTIVTQRRGLMA--LLCVWALSLVISIGPLFG 164

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA 235
              | : ||               ||:     :|:  |:  | ||| :|||:||   ||
Db     165 ---WRQPAP----------EDETICQINEEPGYVLFSALGSFYLPLAIILVMYC---RVY 208

Qy     236 RVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAP---------QTTPH------RTF 280
            ||  :   | : ::| :  || :: |   :    ||          :|  |       :
Db     209 VVAKRESRGLKSGLKTDKSDSEQVTLR---IHRKNAPAGGSGMASAKTKTHFSVRLLKFS 265

Qy     281 GGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSA-----QPISTGQVESVVTWIGYFCFT 335
              |||  |  | | |:|||||:|   | : : :     :|  |  |  :| |:||
Db     266 REKKAAKTLGIVVGCFVLCWLPFF---LVMPIGSFFPDFKPSET--VFKIVFWLGYLNSC 320

Qy     336 SNPFFYGCLNRQIRGELSKQFVCFFKPAPEEEELRL 370
           ||  | | :::           || | :  ||:
Db     321 INPIIYPCSSQE-----------FKKAFQNVLRI 343


Search completed: December 10, 2002, 09:58:24
Job time : 37 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: December 10, 2002, 09:56:20 ; Search time 21 Seconds
(without alignments)
2064.601 Million cell updates/sec

Title: US-09-838-028-2
Perfect score: 2318
Sequence: 1 MESSPIPQSSGNSSTLGRVP.....SDIIMSDSYLRPAASPRLES 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	823	35.5	428	2	I51087	G protein-coupled
2	351	15.1	358	2	JQ1278	histamine H2 recep
3	349.5	15.1	440	2	JC5520	serotonin receptor
4	349	15.1	359	2	A39008	histamine H2 recep
5	343	14.8	359	2	JC4120	histamine H2 recep
6	340	14.7	359	2	JH0449	histamine H2 recep
7	338.5	14.6	436	2	JN0591	serotonin receptor
8	337	14.5	400	2	A53281	beta 3-adrenergic
9	336	14.5	400	2	A41679	beta-3-adrenergic
10	335.5	14.5	405	2	S65459	beta-3-adrenergic
11	335	14.5	428	2	A55044	beta-4C-adrenergic
12	332.5	14.3	572	2	I39369	alpha-1A-adrenergi
13	329.5	14.2	437	2	I57942	5-hydroxytryptamin
14	326.5	14.1	400	2	S32804	beta-3-adrenergic
15	324	14.0	414	1	QRHUB3	beta-3-adrenergic
16	322.5	13.9	418	2	G02953	beta-3-adrenergic
17	322.5	13.9	560	2	A38731	alpha-1A adrenergi
18	320	13.8	501	2	JH0447	alpha-1A-adrenergi
19	319.5	13.8	408	1	QRHUBE	beta-3-adrenergic
20	317.5	13.7	379	2	JC6178	serotonin receptor
21	308	13.3	517	2	A45121	alpha-1B adrenergi
22	305	13.2	515	2	A40491	alpha-1-adrenergic
23	300	12.9	429	2	S65656	alpha-1C-adrenergi
24	300	12.9	466	2	JN0765	alpha-1C-adrenergi
25	300	12.9	499	2	S65657	alpha-1C-adrenergi
26	292.5	12.6	564	2	A38271	serotonin receptor
27	292	12.6	466	2	I57959	alpha-1C adrenergi
28	291.5	12.6	515	2	JC1525	alpha-1B-adrenergi
29	291	12.6	477	2	S71323	alpha-1A adrenergi
30	290.5	12.5	501	2	T18863	hypothetical prote
31	290	12.5	466	2	A35375	alpha-1-adrenergic
32	289.5	12.5	388	2	JN0605	somatostatin recep
33	285	12.3	446	1	DYHUD1	dopamine receptor
34	284.5	12.3	385	2	S68780	dopamine D1-like r
35	283	12.2	446	2	I47217	dopamine receptor
36	281.5	12.1	384	2	A47249	brain-specific som
37	278	12.0	459	2	A56849	dopamine receptor-
38	277.5	12.0	422	2	I38209	serotonin receptor
39	277.5	12.0	483	2	A25896	beta-adrenergic re
40	277.5	12.0	511	2	S44275	dopamine receptor
41	274	11.8	387	2	I49246	D4 dopamine recept
42	272	11.7	444	2	C55886	dopamine receptor
43	271.5	11.7	448	2	A47519	serotonin receptor
44	270.5	11.7	384	2	JC4629	somatostatin recep
45	270	11.6	402	2	I56595	neurokinin 2 recep

ALIGNMENTS

RESULT 1
I51087
G protein-coupled seven-transmembrane receptor - Japanese medaka
C;Species: Oryzias latipes (Japanese medaka)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: I51087
R;Yasuoka, A.; Abe, K.; Saigo, K.; Arai, S.; Emori, Y.
Biochim. Biophys. Acta 1235, 467-469, 1995
A;Title: Molecular cloning of a fish gene encoding a novel seven-transmembrane recept
A;Reference number: I51087; MUID:95275927; PMID:7756357
A;Accession: I51087
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-428 <YAS>
A;Cross-references: GB:D43633; NID:g992581; PIDN:BAA07741.1; PID:g992582
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor

Query Match 35.5%; Score 823; DB 2; Length 428;
Best Local Similarity 42.3%; Pred. No. 3.5e-59;
Matches 184; Conservative 81; Mismatches 142; Indels 28; Gaps 11;

```
Qy       2 ESSPI---PQSSGNSST--LGRVPQTPGPSTASGVPEVGL----RDVASESVALFFMLLL 52
           ::||:     |  | ||   |  |  |        |:||:    :   :   || |: |
Db       5 KTSPMITSDHSISNFSTGLFGPHPTVP--------PDVGVVTSSQSQMKDLFGLFCMVTL 56

Qy      53 DLTAVAGNAAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGE 112
           :| |:  |  ||  ||: | |:|| || ||| ||:| |: |||| ::|||  |   :|
Db      57 NLIALLANTGVMVAIARAPHLKKFAFVCHLCAVDVLCAILLMPLGIISSSPFFGTVVFTI 116

Qy     113 VACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALA 172
           : |::|:||:|  : |:||:::||:||||:|:|||||||:||: ||  |:: :| |:|
Db     117 LECQVYIFLNVFLIWLSILTITAISVERYFYIVHPMRYEVKMTINLVIGVMLLIWFKSLL 176

Qy     173 MASVPVLGRVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMF 232
           :| | : |   :   :      |||  |||    :| |:| |: || |:::|  || :::
Db     177 LALVTLFGWPPYGHQSSIAASHCSLHASHSRLRGVFAVLFCVICFLAPVVVIFSVYSAVY 236

Qy     233 RVARVAAMQHGP-LPTWMET--PRQRSESLSSRSTMVTSSGAPQ-TTPHRTFGGGKAAVV 288
           :||| ||:|  | :||| :    : ||:|::|::|::|:   ||  :| | | |||||:
Db     237 KVARSAALQQVPAVPTWADASPAKDRSDSINSQTTIITTRTLPQRLSPERAFSGGKAALT 296

Qy     289 LLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQI 348
           |  : ||||:||||:| ||| ::|:    | | :|  | |:| |  |  || ||| ||||
Db     297 LAFIVGQFLVCWLPFFIFHLQMSLTGSMKSPGDLEEAVNWLAYSSFAVNPSFYGLLNRQI 356

Qy     349 RGELSK-QFVCFFKPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPK-QE 406
           | || | :  |  :|    |:   | ||| :||||||:| |   ||:  |   :|:  |
Db     357 RDELVKFRRCCVTQPV---EIGPSSLEGSFQENFLQFIQRTSSSSETHPSFANSNPRNME 413

Qy     407 PPAVDFRIPGQIAEE 421
             |   :||||| ||
```

```
Db    414 NQA--HKIPGQIPEE 426


RESULT 2
JQ1278
histamine H2 receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 05-Nov-1999
C;Accession: JQ1278
R;Ruat, M.; Traiffort, E.; Arrang, J.M.; Leurs, R.; Schwartz, J.C.
Biochem. Biophys. Res. Commun. 179, 1470-1478, 1991
A;Title: Cloning and tissue expression of a rat histamine H2-receptor gene.
A;Reference number: JQ1278; MUID:92028890; PMID:1930188
A;Accession: JQ1278
A;Molecule type: DNA
A;Residues: 1-358 <RUA>
A;Cross-references: GB:S57565; NID:g236183; PIDN:AAB19935.1; PID:g236184
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
F;22-45/Domain: transmembrane #status predicted <TM1>
F;58-81/Domain: transmembrane #status predicted <TM2>
F;93-113/Domain: transmembrane #status predicted <TM3>
F;136-159/Domain: transmembrane #status predicted <TM4>
F;178-203/Domain: transmembrane #status predicted <TM5>
F;234-260/Domain: transmembrane #status predicted <TM6>
F;267-288/Domain: transmembrane #status predicted <TM7>
F;4/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;220,311,315/Binding site: phosphate (Ser) (covalent) #status predicted

  Query Match             15.1%;  Score 351;  DB 2;  Length 358;
  Best Local Similarity   27.5%;  Pred. No. 6.2e-21;
  Matches  106;  Conservative   58;  Mismatches  150;  Indels   72;  Gaps   12;

Qy      43 SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLS 100
           || |  ::|:    :|||  |   ::    ||     |:  |   |||  | ::|
Db      22 SVVLTTLILI---TIAGNVVVCLAVSLNRRLRSLTNCFIVSLAATDLLLGLLVLPF---- 74

Qy     101 SSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLG 157
            ||::  :    || | | :|  | |   : :||::  |:::||  |  |:|| | :|
Db      75 -SAIYQLSFTWSFGHVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLVTPV 133

Qy     158 LVASVLVGVWVKALAMASVPVLGRVSWEE------GAPSVPPGCSLQWSHSAYCQLFVVV 211
            ||  || :|| :: :: : :   : |       |  :    | :| :     ::: :|
Db     134 RVAISLVFIWVISITLSFLSI--HLGWNSRNGTRGGNDTFK--CKVQVN-----EVYGLV 184

Qy     212 FAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGA 271
             :: | ||||:: | |  :|::||  | :   : :|
Db     185 DGLVTFYLPLLIMCVTYYRIFKIAREQAKRINHISSWKAA------------------- 224

Qy     272 PQTTPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGY 331
                |    || | | || | |::|| |||:  :|  |          || :| |:||
Db     225 -------TIREHKATVTLAAVMGAFIICWFPYFTAFVYRGLRGDDAINEAVEGIVLWLGY 277

Qy     332 FCFTSNPFFYGCLNRQIRGELSKQFVCFFKP--APEEELRL--------PSREGSIEENF 381
             ||  |  |||  |    : | | |    : :  |||         ||||  :|
Db     278 ANSALNPILYAALNRDFRTAYQQLFHCKFASHNSHKTSLRLNNSLLPRSQSREGRWQEEK 337

Qy     382 LQFLQGTGCPSESWVSRPLPSPKQEP 407
           ||        |    |  |:  |
Db     338 PLKLQ-------VWSGTELTHPQGNP 356


RESULT 3
JC5520
serotonin receptor 6 - human
N;Alternate names: 5-hydroxytryptamine receptor 6 (5-HTR6)
C;Species: Homo sapiens (man)
C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 05-Nov-1999
C;Accession: JC5520
R;Kohen, R.; Metcalf, M.A.; Khan, N.; Druck, T.; Huebner, K.; Lachowicz, J.E.; Meltzer,
J. Neurochem. 66, 47-56, 1996
A;Title: Cloning, characterization, and chromosomal localization of a human 5-HT 6 serot
A;Reference number: JC5520; MUID:96102917; PMID:8522988
A;Accession: JC5520
A;Molecule type: mRNA
A;Residues: 1-440 <KOH>
A;Cross-references: GB:L41147; NID:g1162923; PIDN:AAA92622.1; PID:g1162924
A;Experimental source: brain
C;Comment: This protein shows high affinity for several therapeutically important ant
C;Genetics:
A;Gene: HTR6
A;Map position: 1p35-36
A;Introns: 238/3; 291/3
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; tran
F;28-51/Domain: transmembrane #status predicted <TM1>
F;65-88/Domain: transmembrane #status predicted <TM2>
F;100-123/Domain: transmembrane #status predicted <TM3>
F;143-166/Domain: transmembrane #status predicted <TM4>
F;185-208/Domain: transmembrane #status predicted <TM5>
F;266-289/Domain: transmembrane #status predicted <TM6>
F;298-321/Domain: transmembrane #status predicted <TM7>
F;10/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             15.1%;  Score 349.5;  DB 2;  Length 440;
  Best Local Similarity   30.7%;  Pred. No. 1e-20;
  Matches  127;  Conservative   59;  Mismatches  171;  Indels   57;  Gaps   18;

Qy      19 VPQTPGPSTASGVPEVGLRDVASES----VALFFMLLLDLTAVAGNAAVMAVIAKTPALR 74
           ||: ||| ||:  |  |    ::       ||    ::: ||| | |: ::|:|   ||||
Db       2 VPE-PGP-TANSTPAWGAGPPSAPGGSGWVAAALCVVIALTA-AANSLLIALICTQPALR 58

Qy      75 KF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILS 132
                |:  |   ||:  | :|| |||  :||:   :     | |:     |   | :||:
Db      59 NTSNFFLVSLFTSDLMVGLVVMPPAML--NALYGRWVLARGLCLLWTAFDVMCCSASILN 116

Qy     133 VSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLGRVSWEE---GAP 189
           :  |:::||  :: |:|||::|||     ::::| |  |   : :|:|  : | |     |
Db     117 LCLISLDRYLLILSPLRYKLRMTPLRALALVLGAWSLAALASFLPLL--LGWHELGHARP 174

Qy     190 SVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWM 249
            ||  | |  |       ||:| : | | ||   |   || :   ||  |:|   | | |
Db     175 PVPGQCRLLASLP-----FVLVASGLTFFLPSGAICFTYCRILLAARKQAVQVASLTTGM 229

Qy     250 ----------ETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLL 298
                     |||   ||  ||       |                ||:: |  : | | :
Db     230 ASQASETLQVPRTPRPGVESADSRRLATKHSRKAL----------KASLTLGILLGMFFV 279

Qy     299 CWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVC 358
           |||:|  ::  |:    || |  : |:||:||   | ||  |    |  :  | :
Db     280 TWLPFFVANIVQAV-CDCISPGLFD-VLTWLGYCNSTMNPIIYPLFMRDFKRALGR---- 333

Qy     359 FFKPAPEEELRLP-SREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVD 411
           | | |    | |  |: |:    |:       |  | : : || |   ||  |
Db     334 -FLPCP----RCPRERQASLASPSLRTSHSGPRPGLS-LQQVLPLPL--PPDSD 379


RESULT 4
A39008
histamine H2 receptor - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 24-Nov-1999
C;Accession: A39008; A41821
R;Gantz, I.; Schaeffer, M.; DelValle, J.; Logsdon, C.; Campbell, V.; Uhler, M.; Yamad
Proc. Natl. Acad. Sci. U.S.A. 88, 429-433, 1991
A;Title: Molecular cloning of a gene encoding the histamine H2 receptor.
A;Reference number: A39008; MUID:91110543; PMID:1703298
A;Accession: A39008
A;Molecule type: DNA
A;Residues: 1-359 <GAN>
A;Cross-references: GB:M32701; NID:g163951; PIDN:AAA85637.1; PID:g163952
A;Note: the nucleotide sequence described in this paper has been revised; the amino a
R;Gantz, I.
submitted to the Protein Sequence Database, March 1992
A;Reference number: A41821
A;Accession: A41821
```

```
A;Molecule type: DNA
A;Residues: 1-359 <GA2>
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match            15.1%;  Score 349;  DB 2;  Length 359;
  Best Local Similarity  28.7%;  Pred. No. 9e-21;
  Matches  100;  Conservative   56;  Mismatches  140;  Indels   52;  Gaps   10;

Qy      43 SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLS 100
           || |  ::|:    :||| |   :     ||     |:  | : |||  | ::|
Db      22 SVVLTVLILI---TIAGNVVVCLAVGLNRRLRSLTNCFIVSLAITDLLLGLLVLPF---- 74

Qy     101 SSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLG 157
            || :  :    ||:| | :|  | |   : :||::  |:::||  |  |:|| | :|
Db      75 -SAFYQLSCRWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPV 133

Qy     158 LVASVLVGVWVKALAMASVPV-LGRVSWEEGAP--SVPPGCSLQWSHSAYCQLFVVVFAV 214
            ||  || :|| :: :: : : ||  |  | :      | | :| :      :: :|  :
Db     134 RVAVSLVLIWVISITLSFLSIHLGWNSRNETSSFNHTIPKCKVQVN-----LVYGLVDGL 188

Qy     215 LYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQT 274
           : | ||||:: : |  :|::||  | :   : :|
Db     189 VTFYLPLLVMCITYYRIFKIARDQAKRIHHMGSWKAA---------------------- 225

Qy     275 TPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCF 334
               | |  || | | || | |::|| |||:  :|  |          |:|| |:||
Db     226 ----TIGEHKATVTLAAVMGAFIICWFPYFTVFVYRGLKGDDAINEAFEAVVLWLGYANS 281

Qy     335 TSNPFFYGCLNRQIRGELSKQFVCFFKPAPE--EELRLPSREGSIEEN 380
             ||  |  |||  |    : | |  :||    :|  | |    :  |
Db     282 ALNPILYATLNRDFRTAYQQLFRC--RPASHNAQETSLRSNSSQLARN 327


RESULT 5
JC4120
histamine H2 receptor - guinea pig
C;Species: Cavia porcellus (guinea pig)
C;Date: 27-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 24-Nov-1999
C;Accession: JC4120
R;Traiffort, E.; Vizuete, M.L.; Tardivel-Lacombe, J.; Souil, E.; Schwartz, J.C.; Ruat, M
Biochem. Biophys. Res. Commun. 211, 570-577, 1995
A;Title: The guinea pig histamine H2 receptor: gene cloning, tissue expression and chrom
A;Reference number: JC4120; MUID:95314628; PMID:7794271
A;Accession: JC4120
A;Molecule type: DNA
A;Residues: 1-359 <TRA>
C;Comment: Histamine, a messenger molecule in cell-to-cell communication, affects its ta
ly defined receptor subtypes named H1, H2 and H3.
C;Superfamily: vertebrate rhodopsin
C;Keywords: glycoprotein; phosphoprotein; receptor; transmembrane protein
F;22-45/Domain: transmembrane #status predicted <TM1>
F;58-81/Domain: transmembrane #status predicted <TM2>
F;93-113/Domain: transmembrane #status predicted <TM3>
F;136-157/Domain: transmembrane #status predicted <TM4>
F;179-204/Domain: transmembrane #status predicted <TM5>
F;235-236/Domain: transmembrane #status predicted <TM6>
F;269-289/Domain: transmembrane #status predicted <TM7>
F;4,162,168/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;221,316/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predict
F;226,357/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predict

  Query Match            14.8%;  Score 343;  DB 2;  Length 359;
  Best Local Similarity  27.0%;  Pred. No. 2.8e-20;
  Matches   99;  Conservative   57;  Mismatches  146;  Indels   64;  Gaps   11;

Qy      49 MLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFD 106
           :::| |  ||||  |   :     ||     |:  | : |||  | ::|      ||::
Db      25 LIILILVTVAGNVVVCLAVGLNRRLRSLTNCFIVSLAVTDLLLGLLVLPF-----SAIYQ 79

Qy     107 HAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVL 163
            :    | :| | :|  | |   : :||::  |:::||  |  |:|| | :|   ||  |
Db      80 LSCKWSFSKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPARVAISL 139

Qy     164 VGVWVKALAMASVPVLGRVSWEEGAPSVPPG-----CSLQWSHSAYCQLFVVVFAVLYFL 218
           | :|| :: :: : :   : |     :          | :| :     ::: :|  :: |
Db     140 VFIWVISITLSFLSI--HLGWNSRNETSKDNDTIVKCKVQVN-----EVYGLVDGLVTFY 192

Qy     219 LPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHR 278
           ||||:: : |  :|::||  | :   : :|
Db     193 LPLLIMCITYFRIFKIAREQARRINHIGSWKAA------------------------- 225

Qy     279 TFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNP 338
           |    || | | || | |::|| |||:  :|  |          | || |:||     ||
Db     226 TIREHKATVTLAAVMGAFIICWFPYFTVFVYRGLKGDDAVNEVFEDVVLWLGYANSALNP 285

Qy     339 FFYGCLNRQIRGELSKQFVCFFKPAPEEELRLP---SREGSIEENFLQFLQGTGCPSESW 395
             |  |||  |    : | |          ||    | | |:  |  | |  :  |    |
Db     286 ILYAALNRDFRTAYHQLFCC----------RLASHNSHETSLRLNNSQ-LNRSQCQEPRW 334

Qy     396 -VSRPL 400
            :||
Db     335 QEDKPL 340


RESULT 6
JH0449
histamine H2 receptor - human
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Nov-1999
C;Accession: JH0449; I52319
R;Gantz, I.; Munzert, G.; Tashiro, T.; Schaeffer, M.; Wang, L.; DelValle, J.; Yamada,
Biochem. Biophys. Res. Commun. 178, 1386-1392, 1991
A;Title: Molecular cloning of the human histamine H2 receptor.
A;Reference number: JH0449; MUID:91337087; PMID:1714721
A;Accession: JH0449
A;Molecule type: DNA
A;Residues: 1-359 <GAN>
A;Cross-references: GB:M64799; NID:g184087; PIDN:AAA58647.1; PID:g184088
R;Nishi, T.; Koike, T.; Oka, T.; Maeda, M.; Futai, M.
Biochem. Biophys. Res. Commun. 210, 616-623, 1995
A;Title: Identification of the promoter region of the human histamine H2-receptor gen
A;Reference number: I52319; MUID:95275318; PMID:7755641
A;Accession: I52319
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-359 <RES>
A;Cross-references: GB:D49783; NID:g728495; PIDN:BAA08618.1; PID:g728496
C;Comment: Histamine is important in the regulation of gastric acid secretion.
C;Genetics:
A;Gene: GDB:HRH2
A;Cross-references: GDB:391259; OMIM:142703
A;Map position: 5pter-5qter
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein
F;23-44/Domain: transmembrane #status predicted <TM1>
F;58-81/Domain: transmembrane #status predicted <TM2>
F;93-114/Domain: transmembrane #status predicted <TM3>
F;135-159/Domain: transmembrane #status predicted <TM4>
F;181-204/Domain: transmembrane #status predicted <TM5>
F;235-258/Domain: transmembrane #status predicted <TM6>
F;268-289/Domain: transmembrane #status predicted <TM7>

  Query Match            14.7%;  Score 340;  DB 2;  Length 359;
  Best Local Similarity  28.0%;  Pred. No. 4.9e-20;
  Matches   97;  Conservative   57;  Mismatches  134;  Indels   58;  Gaps    9;

Qy      49 MLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFD 106
           : :| |  |||| |   :     ||     |:  | : |||  | ::|      ||::
Db      25 LAVLILITVAGNVVVCLAVGLNRRLRNLTNCFIVSLAITDLLLGLLVLPF-----SAIYQ 79

Qy     107 HAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVL 163
           :    ||:| | :|  | |   : :||::  |:::||  |: |:|| | :|   ||  |
Db      80 LSCKWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVMDPLRYPVLVTPVRVAISL 139

Qy     164 VGVWVKALAMASVPV-LGRVSWEEGAPS--VPPGCSLQWSHSAYCQLFVVVFAVLYFLLP 220
```

```
Db    140 VLIWVISITLSFLSIHLGWNSRNETSKGNHTTSKCKVQVN-----EVYGLVDGLVTFYLP 194

Qy    221 LLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTF 280
          ||:: : |  :|:|||  | :  : :|                                 |
Db    195 LLIMCITYYRIFKVARDQAKRINHISSWKAA-------------------------TI 227

Qy    281 GGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFF 340
             || | | || | |::|| |||:  :|  |          :|::| |:||     ||
Db    228 REHKATVTLAAVMGAFIICWFPYFTAFVYRGLRGDDAINEVLEAIVLWLGYANSALNPIL 287

Qy    341 YGCLNRQIRGELSKQFVCFFKPAPEEELRLPSREG---SIEENFLQ 383
          |  |||  |    : | |          || :|     |:  |  |
Db    288 YAALNRDFRTGYQQLFCC----------RLANRNSHKTSLRSNASQ 323


RESULT 7
JN0591
serotonin receptor 6 - rat
N;Alternate names: 5-hydroxytryptamine receptor 6 (5-HTR6)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
C;Accession: JN0591
R;Ruat, M.; Traiffort, E.; Arrang, J.M.; Tardivel-Lacombe, J.; Diaz, J.; Leurs, R.; Schw
Biochem. Biophys. Res. Commun. 193, 268-276, 1993
A;Title: A novel rat serotonin (5-HT6) receptor: molecular cloning, localization and sti
A;Reference number: JN0591; MUID:93277562; PMID:8389146
A;Accession: JN0591
A;Molecule type: DNA
A;Residues: 1-436 <RUA>
A;Cross-references: GB:S62043; NID:g385708; PIDN:AAB26908.1; PID:g385709
C;Genetics:
A;Introns: 238/3
C;Superfamily: vertebrate rhodopsin
C;Keywords: glycoprotein; neurotransmitter receptor; transmembrane protein
F;29-53/Domain: transmembrane #status predicted <TM1>
F;63-84/Domain: transmembrane #status predicted <TM2>
F;95-122/Domain: transmembrane #status predicted <TM3>
F;141-168/Domain: transmembrane #status predicted <TM4>
F;185-213/Domain: transmembrane #status predicted <TM5>
F;267-293/Domain: transmembrane #status predicted <TM6>
F;297-319/Domain: transmembrane #status predicted <TM7>
F;9/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             14.6%;  Score 338.5;  DB 2;  Length 436;
  Best Local Similarity   28.4%;  Pred. No. 8e-20;
  Matches  129;  Conservative   58;  Mismatches  171;  Indels   97;  Gaps   20;

Qy      2 ESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNA 61
          |  |:     ||||       |||  | |           |  ||    ::: ||| | |:
Db      4 EPGPV-----NSST---PAWGPGPPPAPG---------GSGWVAAALCVVIVLTA-AANS 45

Qy     62 AVMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYL 119
            :: :|   ||||     |:  |   ||:  | :|| ||| :||:   :      | |:
Db     46 LLIVLICTQPALRNTSNFFLVSLFTSDLMVGLVVMPPAML--NALYGRWVLARGLCLLWT 103

Qy    120 FLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVL 179
             |   | :||::  |:::||  :: |:|||::|||     ::::| |  |   : :|:|
Db    104 AFDVMCCSASILNLCLISLDRYLLILSPLRYKLRMTAPRALALILGAWSLAALASFLPLL 163

Qy    180 GRVSWEE-GAPSVP-PGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMF 232
            : | | |    | ||           |:|     ||:| : : | ||   |    || :
Db    164 --LGWHELGKARTPAPG---------QCRLLASLPFVLVASGVTFFLPSGAICFTYCRIL 212

Qy    233 RVARVAAMQHGPLPTWM-----------ETPRQRSESLSSRSTMVTSSGAPQTTPHRTFG 281
            ||  |:|   | |              |||   ||  ||      |
Db    213 LAARKQAVQVASLTTGTAGQALETLQVPRTPRPGMESADSRRLATKHSRKAL-------- 264

Qy    282 GGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFY 341
            ||:: |  : | | : |||:|  ::  |:    || |  : |:||:||   | ||  |
Db    265 --KASLTLGILLGMFFVTWLPFFVANIAQAV-CDCISPGLFD-VLTWLGYCNSTMNPIIY 320

Qy    342 GCLNRQIRGELSKQFVCFFKPAPEEELRLP----------------------SREGSI 377

Db    321 PLFMRDFKRALGRFLPCVHCP-PEHRPALPPPPCGPLTAVPDQASACSRCCLCLCRQTQI 379

Qy    378 EENFLQFLQGT--GC---PSESWVSRPLPSPKQEP 407
          :      |||    |   ||   : ||  :|:  |
Db    380 QTP----LQGAPRACSSQPSFCCLERPPGTPRHPP 410


RESULT 8
A53281
beta 3-adrenergic receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999
C;Accession: A53281; S29808
R;Granneman, J.G.; Lahners, K.N.; Chaudhry, A.
Mol. Pharmacol. 40, 895-899, 1991
A;Title: Molecular cloning and expression of the rat beta 3-adrenergic receptor.
A;Reference number: A53281; MUID:92100057; PMID:1684635
A;Accession: A53281
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-400 <GRA>
A;Cross-references: GB:S73473; NID:g241215; PIDN:AAB20702.1; PID:g241216
A;Note: sequence extracted from NCBI backbone (NCBIN:73473, NCBIP:73476)
R;Bensaid, M.; Kaghad, M.; Rodriguez, M.; le Fur, G.; Caput, D.
FEBS Lett. 318, 223-226, 1993
A;Title: The rat beta3-adrenergic receptor gene contains an intron.
A;Reference number: S29808; MUID:93178631; PMID:8382630
A;Accession: S29808
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-400 <BEN>
A;Cross-references: GB:S56481; NID:g298306; PIDN:AAB25520.1; PID:g298307
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             14.5%;  Score 337;  DB 2;  Length 400;
  Best Local Similarity   28.2%;  Pred. No. 9.6e-20;
  Matches  120;  Conservative   65;  Mismatches  188;  Indels   52;  Gaps   17;

Qy      4 SPIPQSSGNSSTLGRVPQ-TPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAA 62
          :| |  :|: :     |   |  :  ||:| |          |     || |  | ||
Db      2 APWPHKNGSLAFWSDAPTLDPSAANTSGLPGV-------PWAAALAGALLALATVGGNLL 54

Qy     63 VMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
          |:  ||:|| |:    |||  |   ||:  | :||    :: ||  |   |   | |:
Db     55 VITAIARTPRLQTITNVFVTSLATADLVVGLLVMPPG--ATLALTGHWPLGATGCELWTS 112

Qy    121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
          : |  |: :| :: |: |:||  | :|:||   :|     : :| ||: :  ::  |::
Db    113 VDVLCVTASIETLCALAVDRYLAVTNPLRYGTLVTKRRARAAVVLVWIVSATVSFAPIMS 172

Qy    181 RVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA 235
          :  |  || :     |     |:   |        : :: : : | ||||:|| ||  :| ||
Db    173 Q-WWRVGADAEAQECH---SNPRCCSFASNMPYALLSSSVSFYLPLLVMLFVYARVFVVA 228

Qy    236 ----RVAAMQHGPLPTWMETPR--QRSESLSSRSTMVTSSGAPQ--TTPHR--TFGGGKA 285
              |:   : |  |   |:||   || | ::  |   | | |      | |    |  :|
Db    229 KRQRRLLRRELGRFPP-EESPRSPSRSPSPATVGTPTASDGVPSCGRRPARLLPLGEHRA 287

Qy    286 AVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLN 345
          |  : | | ||||||:|  ::  ||    :    | :|:||      ||  | | :
Db    288 LRTLGLIMGIFSLCWLPFFLANVLRALVGPSLVPSGVFIALNWLGYANSAFNPLIY-CRS 346

Qy    346 RQIRGELSKQFVCFFKPAPEEELR--------LPSREGSIEENFLQFLQGTGCPSESWVS 397
          | :  :: :| :     || |         : ||: |    |       |   |
Db    347 PDFR-DAFRRLLCSYGGRGPEEPRVVTFPASPVASRQNSPLNRF------DGYEGE---- 395

Qy    398 RPLPS 402
          || |:
Db    396 RPFPT 400
```

```
RESULT 9
A41679
beta-3-adrenergic receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 13-Aug-1999
C;Accession: A41679
R;Muzzin, P.; Revelli, J.P.; Kuhne, F.; Gocayne, J.D.; McCombie, W.R.; Venter, J.C.; Gia
J. Biol. Chem. 266, 24053-24058, 1991
A;Title: An adipose tissue-specific beta-adrenergic receptor. Molecular cloning and down
A;Reference number: A41679; MUID:92084710; PMID:1721063
A;Accession: A41679
A;Molecule type: mRNA
A;Residues: 1-400 <MUZ>
A;Cross-references: GB:M74716; NID:g950087; PIDN:AAA74470.1; PID:g202766
A;Experimental source: brown adipose tissue
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F;8,26/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             14.5%;  Score 336;  DB 2;  Length 400;
  Best Local Similarity   28.2%;  Pred. No. 1.2e-19;
  Matches  120;  Conservative   64;  Mismatches  189;  Indels   52;  Gaps   17;

Qy       4 SPIPQSSGNSSTLGRVPQ-TPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAA 62
Db       2 APWPHKNGSLAFWSDAPTLDPSAANTSGLPGV-------PWAAALAGALLALATVGGNLL 54

Qy      63 VMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
Db      55 VITAIARTPRLQTITNVFVTSLATADLVVGLLVMPPG--ATLALTGHWPLGATGCELWTS 112

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
Db     113 VDVLCVTASIETLCALAVDRYLAVTNPLRYGTLVTKRRARAAVVLVWIVSATVSFAPIMS 172

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA 235
Db     173 Q-WWRVGADAEAQECH---SNPRCCSFASNMPYALLSSSVSFYLPLLVMLFVYARVFVVA 228

Qy     236 ----RVAAMQHGPLPTWMETPR--QRSESLSSRSTMVTSSGAPQ--TTPHR--TFGGGKA 285
Db     229 KRQRRFVRRELGRFPP-EESPRSPSRSPSPATVGTPTASDGVPSCGRRPARLLPLGEHRA 287

Qy     286 AVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLN 345
Db     288 LRTLGLIMGIFSLCWLPFFLANVLRALVGPSLVPSGVFIALNWLGYANSAFNPLIY-CRS 346

Qy     346 RQIRGELSKQFVCFFKPAPEEELR--------LPSREGSIEENFLQFLQGTGCPSESWVS 397
Db     347 PDFR-DAFRRLLCSYGGRGPEEPRVVTFPASPVASRQNSPLNRF------DGYEGE---- 395

Qy     398 RPLPS 402
Db     396 RPFPT 400


RESULT 10
S65459
beta-3-adrenergic receptor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 13-Aug-1999
C;Accession: S65459; S40504; S31659
R;Pietri-Rouxel, F.; Lenzen, G.; Kapoor, A.; Drumare, M.F.; Archimbault, P.; Strosberg,
Eur. J. Biochem. 230, 350-358, 1995
A;Title: Molecular cloning and pharmacological characterization of the bovine beta3-adre
A;Reference number: S65459; MUID:95324546; PMID:7601122
A;Accession: S65459
A;Molecule type: mRNA
A;Residues: 1-405 <PIE>
A;Cross-references: EMBL:X85961; NID:g757758; PIDN:CAA59937.1; PID:g757759
A;Experimental source: brown adipose tissue
R;Casteilla, L.; Muzzin, P.; Revelli, J.P.; Ricquier, D.; Giacobino, J.P.
Biochem. J. 297, 93-97, 1994
A;Title: Expression of beta(1)- and beta(3)-adrenergic-receptor messages and adenylat
ite fat.
A;Reference number: S40503; MUID:94107292; PMID:7904157
A;Accession: S40504
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 162-298 <CAS>
R;Stoffel, B.; Meyer, H.H.D.
submitted to the EMBL Data Library, June 1992
A;Description: Bovine beta3-adrenergic receptor, partial genomic sequence.
A;Reference number: S31659
A;Accession: S31659
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 4-5,'HE',8-105 <STO>
A;Cross-references: EMBL:X67214
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; tran
F;37-63/Domain: transmembrane #status predicted <TM1>
F;73-101/Domain: transmembrane #status predicted <TM2>
F;113-133/Domain: transmembrane #status predicted <TM3>
F;156-179/Domain: transmembrane #status predicted <TM4>
F;204-225/Domain: transmembrane #status predicted <TM5>
F;293-314/Domain: transmembrane #status predicted <TM6>
F;327-347/Domain: transmembrane #status predicted <TM7>
F;8,26/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             14.5%;  Score 335.5;  DB 2;  Length 405;
  Best Local Similarity   28.1%;  Pred. No. 1.3e-19;
  Matches  121;  Conservative   67;  Mismatches  183;  Indels   59;  Gaps   18;

Qy       7 PQSSGNSSTLGRVPQTP--GPST--ASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAA 62
Db       3 PWPPGNSS-LTPWPDIPTLAPNTANASGLPGVPW----AVALAGALLALAVLATVGGNLL 57

Qy      63 VMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
Db      58 VIVAIARTPRLQTMTNVFVTSLATADLVVGLLVVPPG--ATLALTGHWPLGVTGCELWTS 115

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
Db     116 VDVLCVTASIETLCALAVDRYLAVTNPLRYGALVTKRRALAAVVLVWVVSAAVSFAPIMS 175

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA 235
Db     176 K-WWRIGADAEAQRCH---SNPRCCTFASNMPYALLSSSVSFYLPLLVMLFVYARVFVVA 231

Qy     236 ---RVAAMQHGPLPTWMETPR-QRSESLSSRSTMVTSSGAP-------QTTPHRTFGGG 283
Db     232 TRQLRLLRRELGRFPPEESPPAPSRSGSPGLAGPCASPAGVPSYGRRPARLLPLREH--- 288

Qy     284 KAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGC 343
Db     289 RALRTLGLIMGTFTLCWLPFFVVNVVRALGGPSLVSGPTFLALNWLGYANSAFNPLIY-C 347

Qy     344 LNRQIRGELSKQFVCFFKPAPEEELRL--PSREGSIEENFLQFLQGTGCPSESWVSRPLP 401
Db     348 RSPDFRSAF-RRLLC--RCRPEEHLAAASPPR---------------APSGAPTALTSP 388

Qy     402 SPKQEPPAVD 411
Db     389 AGPMQPPELD 398


RESULT 11
A55044
beta-4C-adrenergic receptor - turkey
C;Species: Meleagris gallopavo (common turkey)
C;Date: 18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999
C;Accession: A55044
R;Chen, X.; Harden, T.K.; Nicholas, R.A.
J. Biol. Chem. 269, 24810-24819, 1994
A;Title: Molecular cloning and characterization of a novel beta-adrenergic receptor.
```

```
A;Reference number: A55044; MUID:95014249; PMID:7929160
A;Accession: A55044
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-428 <CHE>
A;Cross-references: GB:U13978; NID:g555881; PIDN:AAA62151.1; PID:g555882
C;Genetics:
A;Introns: 416/2
C;Superfamily: vertebrate rhodopsin
C;Keywords: neurotransmitter receptor; transmembrane protein

  Query Match            14.5%;  Score 335;  DB 2;  Length 428;
  Best Local Similarity   27.3%;  Pred. No. 1.5e-19;
  Matches  115;  Conservative   59;  Mismatches  179;  Indels   68;  Gaps   15;

Qy      22 TPGPSTASGVPEVGLRDVASESVALFFMLLLD-LTAVAGNAAVMAVIAKTPALRKF--VF 78
           || |:    ||       | |   |:   | :  |  ||||  |:  ||||| |:    ||
Db       2 TPLPAGNGSVPNCSWAAVLSRQWAVGAALSITILVIVAGNLLVIVAIAKTPRLQTMTNVF 61

Qy      79 VFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILSVSAINV 138
           |  |   ||:  | ::|    ::  |  |  :| | | |:  | |  |: :| :: || |
Db      62 VTSLACADLVMGLLVVPPG--ATILLSGHWPYGTVVCELWTSLDVLCVTASIETLCAIAV 119

Qy     139 ERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLGRVSWEEGAPSVPPGCSLQ 198
           :||  :  |::||  :| |   :|:  ||  :  :: :|::    | :||      |
Db     120 DRYLAITAPLQYEALVTKGRAWAVVCMVWAISAFISFLPIMNH-WWRDGADEQAVRC--- 175

Qy     199 WSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA----------RVAAMQHG 243
           :    |       : :| : : | :|||::: ||  :| ||           :|  :|
Db     176 YDDPRCCDFVTNMTYAIVSSTVSFYVPLLVMIFVYVRVFAVATRHVQLIGKDKVRFLQEN 235

Qy     244 PLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFG--GGKAAVVLLAVGGQFLLCWL 301
           |              |||||       |     | |       ||   |  : | | ||||
Db     236 P------------SLSSR-------GGRWRRPSRLLAIKEHKALKTLGIIMGTFTLCWL 275

Qy     302 PYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFFK 361
           |:|  :: : :  :|:   |:   : |:||     ||  | | :   |    |    |
Db     276 PFFVANI-IKVFCRPLVPDQLFLFLNWLGYVNSAFNPIIY-CRSPDFRSAFRKLLCC--- 330

Qy     362 PAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAEE 421
                | |         | |   |       |   | ::  |||   || : |:
Db     331 ---------PRRADRRLHAAPQDPQHCSC-----AFSPRGDPMEDSKAVD---PGHLRED 373

Qy     422 T 422
           :
Db     374 S 374


RESULT 12
I39369
alpha-1A-adrenergic receptor - human
C;Species: Homo sapiens (man)
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 13-Aug-1999
C;Accession: I39369; JC2331
R;Esbenshade, T.A.; Hirasawa, A.; Tsujimoto, G.; Tanaka, T.; Yano, J.; Minneman, K.P.; M
Mol. Pharmacol. 47, 977-985, 1995
A;Title: Cloning of the human alpha 1d-adrenergic receptor and inducible expression of t
A;Reference number: I39369; MUID:95265059; PMID:7746284
A;Accession: I39369
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-572 <RES>
A;Cross-references: GB:D29952; NID:g914933; PIDN:BAA06222.1; PID:g914934
R;Weinberg, D.H.; Trivedi, P.; Tan, C.P.; Mitra, S.; Perkins-Barrow, A.; Borkowski, D.;
Biochem. Biophys. Res. Commun. 201, 1296-1304, 1994
A;Title: Cloning, expression and characterization of human alpha adrenergic receptors al
A;Reference number: JC2331; MUID:94296402; PMID:8024574
A;Accession: JC2331
A;Molecule type: mRNA
A;Residues: 1-30,'G',32-521,'P',523-572 <WEI>
A;Note: the authors translated the codon CCC for residue 522 as Arg
C;Genetics:
A;Gene: GDB:ADRA1A; ADRA1; ADRA1R
A;Cross-references: GDB:118749; OMIM:104219
A;Map position: 20pter-20qter
A;Introns: 37/3
C;Superfamily: vertebrate rhodopsin
C;Keywords: glycoprotein; neurotransmitter receptor; transmembrane protein
F;89-123/Domain: transmembrane #status predicted <TM1>
F;132-162/Domain: transmembrane #status predicted <TM2>
F;172-197/Domain: transmembrane #status predicted <TM3>
F;208-235/Domain: transmembrane #status predicted <TM4>
F;254-278/Domain: transmembrane #status predicted <TM5>
F;344-374/Domain: transmembrane #status predicted <TM6>
F;380-414/Domain: transmembrane #status predicted <TM7>
F;65,82/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match            14.3%;  Score 332.5;  DB 2;  Length 572;
  Best Local Similarity   24.2%;  Pred. No. 3.3e-19;
  Matches  122;  Conservative   77;  Mismatches  204;  Indels  101;  Gaps   16;

Qy      12 NSSTLGRVPQTPGPSTASGVPEVGLRDVASESVAL-FFMLLLDLTAVAGNAAVMAVIAKT 70
           | |: |           :|   ||   |::: | :  |:    | |||||  |:  :|
Db      65 NRSSAGEPGSAGAGGDVNGTAAVGGLVVSAQGVGVGVFLAAFILMAVAGNLLVILSVACN 124

Qy      71 PALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSL 128
             |:     |: :| : ||| : |::| :  ::  :     ||   | ::  : |   :
Db     125 RHLQTVTNYFIVNLAVADLLLSATVLPFS--ATMEVLGFWAFGRAFCDVWAAVDVLCCTA 182

Qy     129 AILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLGRVSWEEGA 188
           :|||:  |:|:||  | | ::|   ||    |::|  :|| || ::  |:||   |:|
Db     183 SILSLCTISVDRYVGVRHSLKYPAIMTERKAAAILALLWVVALVVSVGPLLG---WKE-- 237

Qy     189 PSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVAR--VAAMQ 241
             |||          :| :     : |  :|  | ||: :|:|:|| :: |||    :::
Db     238 -PVPP-------DERFCGITEEAGYAVFSSVCSFYLPMAVIVVMYCRVYVVARSTTRSLE 289

Qy     242 HGPLPTWMETPRQRSES------LSSRSTMVTSSGA--PQTTPHRTFGGG---------- 283
             |         |:| ::      :  |     : ||   ::    ||
Db     290 AG-------VKRERGKASEVVLRIHCRGAATGADGAHGMRSAKGHTFRSSLSVRLLKFSR 342

Qy     284 --KAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFY 341
             |||  |  | | |:||| |:|      :|  |  :  |  |: |:|||    ||  |
Db     343 EKKAAKTLAIVVGVFVLCWFPFFFVLPLGSLFPQLKPSEGVFKVIFWLGYFNSCVNPLIY 402

Qy     342 GCLNRQIRGELSKQFVCFFK---------------PAPEEELRLPSREGSIEENFLQFL 385
             | :|: :    :   |  :                |   ||      |
Db     403 PCSSREFKRAFLRLLRCQCRRRRRRRPLWRVYGHHWRASTSGLRQDCAPSS--------- 453

Qy     386 QGTGCPSESWVSRPLPSPKQEPPAV-----------------DFRIPGQIAEETSEFLE 427
             |   |        || |  |||                    ::|: |     |:
Db     454 -GDAPPGAPLALTALPDPDPEPPGTPEMQAPVASRRKPPSAFREWRLLGPFRRPTT---- 508

Qy     428 QQLTSDIIMSDSYLRPAASPRLES 451
           || : :      :|   : | |:
Db     509 -QLRAKVSSLSHKIRAGGAQRAEA 531


RESULT 13
I57942
5-hydroxytryptamine receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C;Accession: I57942
R;Monsma, F.J.
Mol. Pharmacol. 43, 320-327, 1993
A;Title: Cloning and expression of a novel serotonin receptor with high affinity for
A;Reference number: I57942; MUID:93196608; PMID:7680751
A;Accession: I57942
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-437 <RES>
A;Cross-references: GB:L03202; NID:g294507; PIDN:AAA40618.1; PID:g294508
C;Superfamily: vertebrate rhodopsin

  Query Match            14.2%;  Score 329.5;  DB 2;  Length 437;
```

```
  Best Local Similarity   27.3%;  Pred. No. 4.3e-19;
  Matches  128;  Conservative   65;  Mismatches  181;  Indels   95;  Gaps   19;

Qy       2 ESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNA 61
Db       4 EPGPV-----NSST---PAWGPGPPPAPG---------GSGWVAAALCVVIVLTA-AANS 45

Qy      62 AVMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYL 119
Db      46 LLIVLICTQPAVRNTSNFFLVSLFTSDLMVGLVVMPPAML--NALYGRWVLARGLCLLWT 103

Qy     120 FLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVL 179
Db     104 AFDVMCCSASILNLCLISLDRYLLILSPLRYKLRMTAPRALALILGAWSLAALASFLPLL 163

Qy     180 GRVSWEE-GAPSVP-PGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMF 232
Db     164 --LGWHELGKARTPAPG---------QCRLLASLPFVLVASGVTFFLPSGAICFTYCRIL 212

Qy     233 RVARVAAMQHGPLPTWM-----------ETPRQRSESLSSRSTMVTSSGAPQTTPHRTFG 281
Db     213 LAARKQAVQVASLTTGTAGQALETLQVPRTPRPGMESADSRRLATKHSRKAL-------- 264

Qy     282 GGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFY 341
Db     265 --KASLTLGILLGMFFVTWLPFFVANIAQAV-CDCISPGLFD-VLTWLGYCNSTMNPIIY 320

Qy     342 GCLNRQIRGELSKQFV----------------------CFFKPAPEEELRLP------S 372
Db     321 PLFMRDFKRALGR-FLHASTVPRSTGQPCLPLHVDLSQRCQTRPQLQQVLALPLPPNSDS 379

Qy     373 REGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAEE 421
Db     380 DSASGGTSGLQLTAQLLLPGEA--TRDPPPPTRATTVVNFFVTDSVEPE 426


RESULT 14
S32804
beta-3-adrenergic receptor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 13-Aug-1999
C;Accession: S32804; S32827; S18548
R;Emorine, L.J.
submitted to the EMBL Data Library, March 1993
A;Reference number: S32803
A;Accession: S32804
A;Molecule type: DNA
A;Residues: 1-400 <EMO>
A;Cross-references: EMBL:X72862; NID:g298112; PIDN:CAA51384.1; PID:g298113
R;van Spronsen, A.; Nahmias, C.; Krief, S.; Briend-Sutren, M.M.; Strosberg, A.D.; Emorin
Eur. J. Biochem. 213, 1117-1124, 1993
A;Title: The promoter and intron/exon structure of the human and mouse beta3-adrenergic
A;Reference number: S32826; MUID:93279311; PMID:8389293
A;Accession: S32827
A;Molecule type: DNA
A;Residues: 1-8;341-347;380-400 <SPR>
R;Nahmias, C.; Blin, N.; Elalouf, J.M.; Mattei, M.G.; Strosberg, A.D.; Emorine, L.J.
EMBO J. 10, 3721-3727, 1991
A;Title: Molecular characterization of the mouse beta-3-adrenergic receptor: relationshi
A;Reference number: S18548; MUID:92037534; PMID:1718744
A;Accession: S18548
A;Molecule type: DNA
A;Residues: 1-388 <NAH>
A;Cross-references: EMBL:X60438; NID:g50109; PIDN:CAA42966.1; PID:g50110
C;Genetics:
A;Map position: 8A2-4
A;Introns: 388/2
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F;37-60/Domain: transmembrane #status predicted <TM1>
F;70-98/Domain: transmembrane #status predicted <TM2>
F;110-130/Domain: transmembrane #status predicted <TM3>
F;153-176/Domain: transmembrane #status predicted <TM4>
F;201-222/Domain: transmembrane #status predicted <TM5>
F;290-311/Domain: transmembrane #status predicted <TM6>
F;324-344/Domain: transmembrane #status predicted <TM7>
F;8,26/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             14.1%;  Score 326.5;  DB 2;  Length 400;
  Best Local Similarity   27.4%;  Pred. No. 6.8e-19;
  Matches  107;  Conservative   61;  Mismatches  171;  Indels   51;  Gaps   12;

Qy       4 SPIPQSSGNSSTLGRVPQ-TPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAA 62
Db       2 APWPHRNGSLALWSDAPTLDPSAANTSGLPGV-------PWAAALAGALLALATVGGNLL 54

Qy      63 VMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
Db      55 VIIAIARTPRLQTITNVFVTSLAAADLVVGLLVMPPG--ATLALTGHWPLGETGCELWTS 112

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
Db     113 VDVLCVTASIETLCALAVDRYLAVTNPLRYGTLVTKRRARAAVVLVWIVSAAVSFAPIMS 172

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA 235
Db     173 Q-WWRVGADAEAQECH---SNPRCCSFASNMPYALLSSSVSFYLPLLVMLFVYARVFVVA 228

Qy     236 ------------RVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAP-------QTTP 276
Db     229 KRQRHLLRRELGRFSPEESPPSPS-------RSPSPATGGTPAAPDGVPPCGRRPARLLP 281

Qy     277 HRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTS 336
Db     282 LREH---RALRTLGLIMGIFSLCWLPFFLANVLRALAGPSLVPSGVFIALNWLGYANSAF 338

Qy     337 NPFFYGCLNRQIRGELSKQFVCFFKPAPEE 366
Db     339 NPVIY-CRSPDFRDAFRRLLCSYGGRGPEE 367


RESULT 15
QRHUB3
beta-3-adrenergic receptor, splice form 2 - human
N;Alternate names: beta-3-adrenergic receptor form A
N;Contains: beta-3-adrenergic receptor splice form 1
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1993 #sequence_revision 18-Aug-1995 #text_change 16-Jun-2000
C;Accession: A41348; S33752
R;Emorine, L.J.; Marullo, S.; Briend-Sutren, M.M.; Patey, G.; Tate, K.; Delavier-Klut
Science 245, 1118-1121, 1989
A;Title: Molecular characterization of the human beta-3-adrenergic receptor.
A;Reference number: A41348; MUID:89368947; PMID:2570461
A;Accession: A41348
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-402 <EMO>
A;Cross-references: GB:M29932; NID:g178895; PIDN:AAA35550.1; PID:g178896
A;Note: splice form 1
R;Lelias, J.M.; Kaghad, M.; Rodriguez, M.; Chalon, P.; Bonnin, J.; Dupre, I.; Delpech
FEBS Lett. 324, 127-130, 1993
A;Title: Molecular cloning of a human beta-3-adrenergic receptor cDNA.
A;Reference number: S33751; MUID:93285320; PMID:8389717
A;Accession: S33752
A;Molecule type: DNA
A;Residues: 392-414 <LEL>
A;Cross-references: EMBL:X70812; NID:g312398; PIDN:CAA50142.1; PID:g1666375
A;Note: splice form 2
C;Genetics:
A;Gene: GDB:ADRB3
A;Cross-references: GDB:203869; OMIM:109691
A;Map position: 8p12-8p11.1
A;Introns: 402/2
C;Superfamily: vertebrate rhodopsin
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmemb
F;1-414/Product: beta-3-adrenergic receptor precursor splice form 2 #status predicted
```

```
F;1-402/Product: beta-3-adrenergic receptor precursor splice form 1 #status predicted <M
F;37-63/Domain: transmembrane #status predicted <TM1>
F;73-101/Domain: transmembrane #status predicted <TM2>
F;113-133/Domain: transmembrane #status predicted <TM3>
F;156-178/Domain: transmembrane #status predicted <TM4>
F;204-225/Domain: transmembrane #status predicted <TM5>
F;293-314/Domain: transmembrane #status predicted <TM6>
F;327-347/Domain: transmembrane #status predicted <TM7>
F;8,26/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             14.0%;  Score 324;  DB 1;  Length 414;
  Best Local Similarity   28.7%;  Pred. No. 1.1e-18;
  Matches  125;  Conservative   63;  Mismatches  187;  Indels   60;  Gaps   21;

Qy       4 SPIPQSSGNSSTLGRVPQTP--GPSTA--SGVPEVGLRDVASESVALFFMLLLDLTAVAG 59
Db       2 APWPH---ENSSLAPWPDLPTLAPNTANTSGLPGVPW----EAALAGALLALAVLATVGG 54

Qy      60 NAAVMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRL 117
Db      55 NLLVIVAIAWTPRLQTMTNVFVTSLAAADLVMGLLVVPPA--ATLALTGHWPLGATGCEL 112

Qy     118 YLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVP 177
Db     113 WTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALVTKRCARTAVVLVWVVSAAVSFAP 172

Qy     178 VLGRVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMF 232
Db     173 IMSQ-WWRVGADAEAQRCH---SNPRCCAFASNMPYVLLSSSVSFYLPLLVMLFVYARVF 228

Qy     233 RVA----RVAAMQHGPLPTWMETPRQRSESLSSR--STMVTSSGAP-------QTTPHRT 279
Db     229 VVATRQLRLLRGELGRFPP-EESPPAPSRSLAPAPVGTCAPPEGVPACGRRPARLLPLRE 287

Qy     280 FGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPF 339
Db     288 H---RALCTLGLIMGTFTLCWLPFFLANVLRALGGPSLVPGPAFLALNWLGYANSAFNPL 344

Qy     340 FYGCLNRQIRGELSKQFVC----FFKPAPEEELR---LPS-----REGSIEENFLQFLQG 387
Db     345 IY-CRSPDFRSAF-RRLLCRCGRRLPPEPCAAARPALFPSGVPAARSSPAQPRLCQRLDG 402

Qy     388 -TGCPSESWVSRPLP 401
Db     403 FDSCGGEN----PLP 413


Search completed: December 10, 2002, 09:58:51
Job time : 22 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: December 10, 2002, 09:56:35 ; Search time 18 Seconds
(without alignments)
737.208 Million cell updates/sec

Title: US-09-838-028-2
Perfect score: 2318
Sequence: 1 MESSPIPQSSGNSSTLGRVP.....SDIIMSDSYLRPAASPRLES 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1276	55.0	252	3	US-08-513-974B-52	Sequence 52, Appl
2	1276	55.0	252	3	US-08-513-974B-360	Sequence 360, App
3	351	15.1	358	2	US-08-748-485-6	Sequence 6, Appli
4	350	15.1	359	2	US-08-103-170-7	Sequence 7, Appli
5	349	15.1	359	1	US-07-996-772A-10	Sequence 10, Appl
6	349	15.1	359	2	US-08-748-485-3	Sequence 3, Appli
7	349	15.1	359	4	US-09-328-314-19	Sequence 19, Appl
8	347	15.0	348	3	US-08-875-540-13	Sequence 13, Appl
9	347	15.0	348	4	US-09-171-456-17	Sequence 17, Appl
10	347	15.0	348	4	US-09-473-634-13	Sequence 13, Appl
11	346	14.9	359	2	US-08-103-170-4	Sequence 4, Appli
12	346	14.9	359	2	US-08-103-170-6	Sequence 6, Appli
13	343	14.8	359	2	US-08-748-485-5	Sequence 5, Appli
14	340	14.7	359	2	US-08-467-568-13	Sequence 13, Appl
15	340	14.7	359	2	US-08-748-485-4	Sequence 4, Appli
16	340	14.7	359	2	US-08-103-170-2	Sequence 2, Appli
17	340	14.7	359	2	US-09-030-582-13	Sequence 13, Appl
18	340	14.7	359	3	US-08-875-540-15	Sequence 15, Appl
19	340	14.7	359	4	US-09-171-456-19	Sequence 19, Appl
20	340	14.7	359	4	US-09-473-634-15	Sequence 15, Appl
21	337	14.5	400	1	US-07-916-901-6	Sequence 6, Appli
22	336	14.5	400	1	US-07-783-602C-1	Sequence 1, Appli
23	335.5	14.5	405	1	US-08-351-473B-2	Sequence 2, Appli
24	335	14.5	400	1	US-08-351-473B-4	Sequence 4, Appli
25	332.5	14.3	572	1	US-08-334-698-2	Sequence 2, Appli
26	332.5	14.3	572	1	US-08-228-932-2	Sequence 2, Appli
27	332.5	14.3	572	1	US-08-468-939-2	Sequence 2, Appli
28	332.5	14.3	572	1	US-08-722-001-30	Sequence 30, Appl
29	332.5	14.3	572	2	US-08-406-855A-2	Sequence 2, Appli
30	332.5	14.3	572	2	US-08-722-190-2	Sequence 2, Appli
31	332.5	14.3	572	3	US-08-244-354-2	Sequence 2, Appli
32	332.5	14.3	572	3	US-09-206-899-2	Sequence 2, Appli
33	332.5	14.3	572	4	US-09-444-783-2	Sequence 2, Appli
34	332.5	14.3	572	4	US-09-688-415-2	Sequence 2, Appli
35	332.5	14.3	572	5	PCT-US95-04203-2	Sequence 2, Appli
36	326.5	14.1	388	1	US-08-087-772A-2	Sequence 2, Appli
37	326.5	14.1	400	1	US-08-351-473B-5	Sequence 5, Appli
38	326.5	14.1	400	4	US-08-450-962-4	Sequence 4, Appli
39	326.5	14.1	400	4	US-08-450-962-6	Sequence 6, Appli
40	325	14.0	501	4	US-09-688-415-7	Sequence 7, Appli
41	322.5	13.9	560	4	US-09-688-415-8	Sequence 8, Appli
42	321	13.8	559	2	US-08-406-855A-20	Sequence 20, Appl
43	321	13.8	559	3	US-09-206-899-20	Sequence 20, Appl
44	320	13.8	501	1	US-08-722-001-14	Sequence 14, Appl
45	320	13.8	501	2	US-08-467-568-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-08-513-974B-52
; Sequence 52, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Ohtaki, Tetsuya
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Ohgi, Kazuhiro
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
; NUMBER OF SEQUENCES: 380
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,974B
; FILING DATE: 14-SEP-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01599
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-093989
; FILING DATE: 19-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-057186
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-007177
; FILING DATE: 20-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-326611
; FILING DATE: 28-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-270017
; FILING DATE: 02-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236357

```
;      FILING DATE:  30-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-236356
;      FILING DATE:  30-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-189274
;      FILING DATE:  11-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-189273
;      FILING DATE:  11-AUG-1945
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-189272
;      FILING DATE:  11-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Resnick, David S.
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  45753
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;  INFORMATION FOR SEQ ID NO:  52:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  252 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-513-974B-52

  Query Match             55.0%;  Score 1276;  DB 3;  Length 252;
  Best Local Similarity   98.0%;  Pred. No. 1.9e-92;
  Matches  247;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      85 VDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILSVSAINVERYYYV 144
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 VDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILSVSAINVERYYYV 60

Qy     145 VHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLGRVSWEEGAPSVPPGCSLQWSHSAY 204
           |||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||
Db      61 VHPMRYEVRMKLGLVASVLVGVWVKALAMASVPVLGRVSWEEGPPSVPPGCSLQWSHSAY 120

Qy     205 CQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRST 264
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRST 180

Qy     265 MVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVES 324
           |||||||||||||||||||||||||||||||||||||||||||||||||||: |||| :
Db     181 MVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPIAAGQVEN 240

Qy     325 VVTWIGYFCFTS 336
           ||||||||||||
Db     241 VVTWIGYFCFTS 252


RESULT 2
US-08-513-974B-360
; Sequence 360, Application US/08513974B
; Patent No. 6114139
;  GENERAL INFORMATION:
;    APPLICANT:  Hinuma, Shuji
;    APPLICANT:  Hosoya, Masaki
;    APPLICANT:  Fujii, Ryo
;    APPLICANT:  Ohtaki, Tetsuya
;    APPLICANT:  Fukusumi, Shoji
;    APPLICANT:  Ohgi, Kazuhiro
;    TITLE OF INVENTION:  G PROTEIN COUPLED RECEPTOR PROTEIN,
;    TITLE OF INVENTION:  PRODUCTION, AND USE THEREOF
;    NUMBER OF SEQUENCES:  380
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 Water Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/513,974B
;      FILING DATE:  14-SEP-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP95/01599
;      FILING DATE:  10-AUG-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-093989
;      FILING DATE:  19-AUG-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-057186
;      FILING DATE:  16-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-007177
;      FILING DATE:  20-JAN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-326611
;      FILING DATE:  28-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-270017
;      FILING DATE:  02-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-236357
;      FILING DATE:  30-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-236356
;      FILING DATE:  30-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-189274
;      FILING DATE:  11-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-189273
;      FILING DATE:  11-AUG-1945
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-189272
;      FILING DATE:  11-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Resnick, David S.
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  45753
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;  INFORMATION FOR SEQ ID NO:  360:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  252 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-513-974B-360

  Query Match             55.0%;  Score 1276;  DB 3;  Length 252;
  Best Local Similarity   98.0%;  Pred. No. 1.9e-92;
  Matches  247;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      85 VDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILSVSAINVERYYYV 144
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 VDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILSVSAINVERYYYV 60

Qy     145 VHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLGRVSWEEGAPSVPPGCSLQWSHSAY 204
           |||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||
Db      61 VHPMRYEVRMKLGLVASVLVGVWVKALAMASVPVLGRVSWEEGPPSVPPGCSLQWSHSAY 120

Qy     205 CQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRST 264
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    121 CQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRST 180

Qy    265 MVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVES 324
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||: ||||:
Db    181 MVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPIAAGQVEN 240

Qy    325 VVTWIGYFCFTS 336
          ||||||||||||
Db    241 VVTWIGYFCFTS 252


RESULT 3
US-08-748-485-6
; Sequence 6, Application US/08748485
; Patent No. 5817480
;  GENERAL INFORMATION:
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Murry, Lynn E.
;    TITLE OF INVENTION:  NOVEL HISTAMINE H2 RECEPTOR
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,485
;      FILING DATE:  Herewith
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0159 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  358 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  236184
US-08-748-485-6

  Query Match           15.1%;  Score 351;  DB 2;  Length 358;
  Best Local Similarity   27.5%;  Pred. No. 5e-20;
  Matches  106;  Conservative   58;  Mismatches  150;  Indels   72;  Gaps   12;

Qy     43 SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLS 100
          || |  ::|:    :||| |   ::    ||     |:  |   |||  | ::|
Db     22 SVVLTTLILI---TIAGNVVVCLAVSLNRRLRSLTNCFIVSLAATDLLLGLLVLPF---- 74

Qy    101 SSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLG 157
           ||::  :    || | | :|  | |   : :||::  |:::||  |  |:|| | :|
Db     75 -SAIYQLSFTWSFGHVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLVTPV 133

Qy    158 LVASVLVGVWVKALAMASVPVLGRVSWEE------GAPSVPPGCSLQWSHSAYCQLFVVV 211
           ||  || :|| :: :: : :   : |        |  :    | :| :     ::: :|
Db    134 RVAISLVFIWVISITLSFLSI--HLGWNSRNGTRGGNDTFK--CKVQVN-----EVYGLV 184

Qy    212 FAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGA 271
           :: | ||||:: | |  :|::||  | :   : :|
Db    185 DGLVTFYLPLLIMCVTYYRIFKIAREQAKRINHISSWKAA------------------- 224

Qy    272 PQTTPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGY 331
                |    || | | || | |::|| |||:  :|  |         || :| |:||
Db    225 -------TIREHKATVTLAAVMGAFIICWFPYFTAFVYRGLRGDDAINEAVEGIVLWLGY 277

Qy    332 FCFTSNPFFYGCLNRQIRGELSKQFVCFFKP--APEEELRL--------PSREGSIEENF 381
             ||  |  |||  |    : | | |    : :  |||         ||||  :|
Db    278 ANSALNPILYAALNRDFRTAYQQLFHCKFASHNSHKTSLRLNNSLLPRSQSREGRWQEEK 337

Qy    382 LQFLQGTGCPSESWVSRPLPSPKQEP 407
             ||        |    |  |:  |
Db    338 PLKLQ-------VWSGTELTHPQGNP 356


RESULT 4
US-08-103-170-7
; Sequence 7, Application US/08103170
; Patent No. 5885824
;  GENERAL INFORMATION:
;    APPLICANT:  Yamada, Tadataka
;    APPLICANT:  Gantz, Ira
;    TITLE OF INVENTION:  Recombinant Genomic Clones Encoding
;    TITLE OF INVENTION:  Histamine H1, H2, and H3 Receptors, Methods For Production
;    TITLE OF INVENTION:  Thereof, and Proteins Encoded Therefrom
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 Jefferson Davis Highway, Fourth Floor
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/103,170
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/633,060
;      FILING DATE:  24-DEC-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lavalleye, Jean-Paul
;      REGISTRATION NUMBER:  31,451
;      REFERENCE/DOCKET NUMBER:  2363-017-55
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)521-4500
;      TELEFAX:  (703)486-2347
;      TELEX:  24885 OPAT UR
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Dog
US-08-103-170-7

  Query Match           15.1%;  Score 350;  DB 2;  Length 359;
  Best Local Similarity   28.7%;  Pred. No. 6e-20;
  Matches  100;  Conservative   56;  Mismatches  140;  Indels   52;  Gaps   10;
```

```
Qy      43 SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLS 100
           || |  ::|:    :|||  |   :     ||     |:  | : |||  | ::|
Db      22 SVVLTVLILI---TIAGNVVVCLAVGLNRRLRSLTNCFIVSLAITDLLLGLLVLPF---- 74

Qy     101 SSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLG 157
            || :  :    ||:| | :|  | |   : :||::  |:::||  |  |:|| | :|
Db      75 -SAFYQLSCRWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPV 133

Qy     158 LVASVLVGVWVKALAMASVPV-LGRVSWEEGAP--SVPPGCSLQWSHSAYCQLFVVVFAV 214
           ||  || :|| :: :: : :|| |  | :      | | :| :      :: :|  :
Db     134 RVAVSLVLIWVISITLSFLSIHLGWNSTNETSSFNHTIPKCKVQVN-----LVYGLVDGL 188

Qy     215 LYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQT 274
           : | ||||:: : |  :|::||  | :   : :|
Db     189 VTFYLPLLVMCITYYRIFKIARDQAKRIHHMGSWKAA---------------------- 225

Qy     275 TPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCF 334
               | |  || | | || | |::|| |||:  :|  |           |:|| |:||
Db     226 ----TIGEHKATVTLAAVMGAFIICWFPYFTVFVYRGLKGDDAINEAFEAVVLWLGYANS 281

Qy     335 TSNPFFYGCLNRQIRGELSKQFVCFFKPAPE--EELRLPSREGSIEEN 380
             ||  |  |||  |    : | |  :||    :|  | |    :  |
Db     282 ALNPILYATLNRDFRTAYQQLFRC--RPASHNAQETSLRSNSSQLARN 327


RESULT 5
US-07-996-772A-10
; Sequence 10, Application US/07996772A
; Patent No. 5472866
;  GENERAL INFORMATION:
;    APPLICANT:  Gerald, Christophe
;    APPLICANT:  Hartig, Paul R.
;    APPLICANT:  Branchek, Theresa A.
;    APPLICANT:  Weinshank, Richard L.
;    TITLE OF INVENTION:  DNA ENCODING 5-HT4A SEROTONIN
;    TITLE OF INVENTION:  RECEPTORS AND USES THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  COOPER & DUNHAM
;      STREET:  30 ROCKEFELLER PLAZA
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      ZIP:  10112
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/996,772A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, P., John
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  42667/JPW/TEP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 977-9550
;      TELEFAX:  (212) 664-0525
;      TELEX:  422523 COOP UI
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    IMMEDIATE SOURCE:
;      CLONE:  HIST2
US-07-996-772A-10

  Query Match           15.1%;  Score 349;  DB 1;  Length 359;
  Best Local Similarity  28.7%;  Pred. No. 7.1e-20;
  Matches  100;  Conservative   56;  Mismatches  140;  Indels   52;  Gaps   10;

Qy      43 SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLS 100
           || |  ::|:    :|||  |   :     ||     |:  | : |||  | ::|
Db      22 SVVLTVLILI---TIAGNVVVCLAVGLNRRLRSLTNCFIVSLAITDLLLGLLVLPF---- 74

Qy     101 SSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLG 157
            || :  :    ||:| | :|  | |   : :||::  |:::||  |  |:|| | :|
Db      75 -SAFYQLSCRWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPV 133

Qy     158 LVASVLVGVWVKALAMASVPV-LGRVSWEEGAP--SVPPGCSLQWSHSAYCQLFVVVFAV 214
           ||  || :|| :: :: : :|| |  | :      | | :| :      :: :|  :
Db     134 RVAVSLVLIWVISITLSFLSIHLGWNSRNETSSFNHTIPKCKVQVN-----LVYGLVDGL 188

Qy     215 LYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQT 274
           : | ||||:: : |  :|::||  | :   : :|
Db     189 VTFYLPLLVMCITYYRIFKIARDQAKRIHHMGSWKAA---------------------- 225

Qy     275 TPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCF 334
               | |  || | | || | |::|| |||:  :|  |           |:|| |:||
Db     226 ----TIGEHKATVTLAAVMGAFIICWFPYFTVFVYRGLKGDDAINEAFEAVVLWLGYANS 281

Qy     335 TSNPFFYGCLNRQIRGELSKQFVCFFKPAPE--EELRLPSREGSIEEN 380
             ||  |  |||  |    : | |  :||    :|  | |    :  |
Db     282 ALNPILYATLNRDFRTAYQQLFRC--RPASHNAQETSLRSNSSQLARN 327


RESULT 6
US-08-748-485-3
; Sequence 3, Application US/08748485
; Patent No. 5817480
;  GENERAL INFORMATION:
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Murry, Lynn E.
;    TITLE OF INVENTION:  NOVEL HISTAMINE H2 RECEPTOR
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,485
;      FILING DATE:  Herewith
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0159 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
```

```
;      CLONE:  163952
US-08-748-485-3

  Query Match             15.1%;  Score 349;  DB 2;  Length 359;
  Best Local Similarity   28.7%;  Pred. No. 7.1e-20;
  Matches  100;  Conservative   56;  Mismatches  140;  Indels   52;  Gaps   10;

Qy      43 SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLS 100
           || |  ::|:    :|||  |   :     ||     |:  | : |||  | ::|
Db      22 SVVLTVLILI---TIAGNVVVCLAVGLNRRLRSLTNCFIVSLAITDLLLGLLVLPF---- 74

Qy     101 SSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLG 157
            || :  :    ||:| | :|  | |   : :||::  |:::||  |  |:|| | :|
Db      75 -SAFYQLSCRWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPV 133

Qy     158 LVASVLVGVWVKALAMASVPV-LGRVSWEEGAP--SVPPGCSLQWSHSAYCQLFVVVFAV 214
           ||  || :|| :: :: : : ||  |  | :      | | :| :      :: :|  :
Db     134 RVAVSLVLIWVISITLSFLSIHLGWNSRNETSSFNHTIPKCKVQVN-----LVYGLVDGL 188

Qy     215 LYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQT 274
           : | ||||:: : |  :|::||  | :   : :|
Db     189 VTFYLPLLVMCITYYRIFKIARDQAKRIHHMGSWKAA---------------------- 225

Qy     275 TPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCF 334
               | |  || | | || | |::|| |||:  :|  |          |:|| |:||
Db     226 ----TIGEHKATVTLAAVMGAFIICWFPYFTVFVYRGLKGDDAINEAFEAVVLWLGYANS 281

Qy     335 TSNPFFYGCLNRQIRGELSKQFVCFFKPAPE--EELRLPSREGSIEEN 380
           ||  |  |||  |    : | |  :||    :|  | |    :  |
Db     282 ALNPILYATLNRDFRTAYQQLFRC--RPASHNAQETSLRSNSSQLARN 327


RESULT 7
US-09-328-314-19
; Sequence 19, Application US/09328314
; Patent No. 6331401
; GENERAL INFORMATION:
;  APPLICANT: Gerald, Christophe
;  APPLICANT: Hartig, Paul R.
;  APPLICANT: Branchek, Theresa
;  APPLICANT: Weinshank, Richard L.
;  TITLE OF INVENTION: DNA Encoding 5-HT4 Serotonin Receptors And Uses Thereof
;  FILE REFERENCE: 42667-AZ-PCT-US
;  CURRENT APPLICATION NUMBER: US/09/328,314
;  CURRENT FILING DATE: 1998-04-03
;  EARLIER APPLICATION NUMBER: 08/446,822
;  EARLIER FILING DATE: 1995-07-31
;  EARLIER APPLICATION NUMBER: PCT/US93/12586
;  EARLIER FILING DATE: 1993-12-22
;  EARLIER APPLICATION NUMBER: 07/996,772
;  EARLIER FILING DATE: 1992-12-24
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: canine
US-09-328-314-19

  Query Match             15.1%;  Score 349;  DB 4;  Length 359;
  Best Local Similarity   28.7%;  Pred. No. 7.1e-20;
  Matches  100;  Conservative   56;  Mismatches  140;  Indels   52;  Gaps   10;

Qy      43 SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLS 100
           || |  ::|:    :|||  |   :     ||     |:  | : |||  | ::|
Db      22 SVVLTVLILI---TIAGNVVVCLAVGLNRRLRSLTNCFIVSLAITDLLLGLLVLPF---- 74

Qy     101 SSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLG 157
            || :  :    ||:| | :|  | |   : :||::  |:::||  |  |:|| | :|
Db      75 -SAFYQLSCRWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPV 133

Qy     158 LVASVLVGVWVKALAMASVPV-LGRVSWEEGAP--SVPPGCSLQWSHSAYCQLFVVVFAV 214
           ||  || :|| :: :: : : ||  |  | :      | | :| :      :: :|  :
Db     134 RVAVSLVLIWVISITLSFLSIHLGWNSRNETSSFNHTIPKCKVQVN-----LVYGLVDGL 188

Qy     215 LYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQT 274
           : | ||||:: : |  :|::||  | :   : :|
Db     189 VTFYLPLLVMCITYYRIFKIARDQAKRIHHMGSWKAA---------------------- 225

Qy     275 TPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCF 334
               | |  || | | || | |::|| |||:  :|  |          |:|| |:||
Db     226 ----TIGEHKATVTLAAVMGAFIICWFPYFTVFVYRGLKGDDAINEAFEAVVLWLGYANS 281

Qy     335 TSNPFFYGCLNRQIRGELSKQFVCFFKPAPE--EELRLPSREGSIEEN 380
           ||  |  |||  |    : | |  :||    :|  | |    :  |
Db     282 ALNPILYATLNRDFRTAYQQLFRC--RPASHNAQETSLRSNSSQLARN 327


RESULT 8
US-08-875-540-13
; Sequence 13, Application US/08875540A
; Patent No. 6015888
; GENERAL INFORMATION:
;  APPLICANT: Heath, Paul Roy
;  APPLICANT: Orange, Paul Richard
;  APPLICANT: Pearson, Ronald Carl Alan
;  APPLICANT: Wright, Simon Ralph
;  TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO THE DETECTION OF VARIATIONS IN
;  TITLE OF INVENTION: HUMAN HISTAMINE H2 RECEPTORS
;  FILE REFERENCE: 09347/002001
;  CURRENT APPLICATION NUMBER: US/08/875,540A
;  CURRENT FILING DATE: 1998-01-05
;  EARLIER APPLICATION NUMBER: PCT/EP96/00397
;  EARLIER FILING DATE: 1997-01-30
;  EARLIER APPLICATION NUMBER: GB9503866.7
;  EARLIER FILING DATE: 1995-01-30
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-875-540-13

  Query Match             15.0%;  Score 347;  DB 3;  Length 348;
  Best Local Similarity   28.6%;  Pred. No. 9.9e-20;
  Matches   99;  Conservative   57;  Mismatches  132;  Indels   58;  Gaps   10;

Qy      49 MLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFD 106
           : :| |  |||| |   :     ||     |:  | : |||  | ::|   '   ||::
Db      22 LAVLILITVAGNVVVCLAVGLNRRLRNLTNCFIVSLAITDLLLGLLVLPF-----SAIYQ 76

Qy     107 HAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVL 163
             :    ||:| | :|  | |   : :||::  |:::||  |: |:|| | :|   ||  |
Db      77 LSCKWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVMDPLRYPVLVTPARVAISL 136

Qy     164 VGVWVKALAMASVPV-LGRVSWEEGAPS--VPPGCSLQWSHSAYCQLFVVVFAVLYFLLP 220
           | :|| :: :: : : ||  |  | :        |::| :     ::: :|  :: | ||
Db     137 VLIWVISITLSFLSIHLGWNSRNETSKGNHTTSKCNVQVN-----EVYGLVDGLVTFYLP 191

Qy     221 LLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTF 280
           ||:: : |  :|||||  | :   : :|                 |   ||
Db     192 LLIMCITYYRIFRVARDQAKRIDHISSWK-------------------AATIREHR-- 228

Qy     281 GGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFF 340
           | | | || | |::|| |||:  :|  |          :|::| |:||     ||
Db     229 ----ATVTLAAVMGAFIICWFPYFTAFVYRGLRGDDAINEMLEAIVLWLGYANSALNPIL 284

Qy     341 YGCLNRQIRGELSKQFVCFFKPAPEEELRLPSREG---SIEENFLQ 383
           |  |||  |    : | |           || :|     |:  |  |
Db     285 YAALNRDFRTGYQQLFCC----------RLANRNSHKTSLRSNASQ 320


RESULT 9
US-09-171-456-17
; Sequence 17, Application US/09171456A
```

```
; Patent No. 6346380
; GENERAL INFORMATION:
;  APPLICANT: Wright, Simon Ralph
;  APPLICANT: Heath, Paul Roy
;  APPLICANT: Orange, Paul Richard
;  APPLICANT: Pearson, Ronald Carl Alan
;  TITLE OF INVENTION: DETECTION OF VARIATIONS IN HUMAN H2 RECEPTORS
;  FILE REFERENCE: 09347/004001
;  CURRENT APPLICATION NUMBER: US/09/171,456A
;  CURRENT FILING DATE: 1999-08-03
;  EARLIER APPLICATION NUMBER: PCT/GB97/01075
;  EARLIER FILING DATE: 1997-04-18
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-171-456-17

  Query Match             15.0%;  Score 347;  DB 4;  Length 348;
  Best Local Similarity   28.6%;  Pred. No. 9.9e-20;
  Matches   99;  Conservative   57;  Mismatches  132;  Indels   58;  Gaps   10;

Qy      49 MLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFD 106
           : :| |  |||| |   :     ||     |:  | : |||  | ::|      ||::
Db      22 LAVLILITVAGNVVVCLAVGLNRRLRNLTNCFIVSLAITDLLLGLLVLPF-----SAIYQ 76

Qy     107 HAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVL 163
            :    ||:| | :|  | |   : :||::  |:::||  |: |:|| | :|   ||  |
Db      77 LSCKWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVMDPLRYPVLVTPARVAISL 136

Qy     164 VGVWVKALAMASVPV-LGRVSWEEGAPS--VPPGCSLQWSHSAYCQLFVVVFAVLYFLLP 220
           | :|| :: :: : : ||  |  | :      |::| :     ::: :|  :: | ||
Db     137 VLIWVISITLSFLSIHLGWNSRNETSKGNHTTSKCNVQVN-----EVYGLVDGLVTFYLP 191

Qy     221 LLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTF 280
           ||:: : |  :|||||  | :   : :|                    |     ||
Db     192 LLIMCITYYRIFRVARDQAKRIDHISSWK--------------------AATIREHR-- 228

Qy     281 GGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFF 340
               | | | || | |::|| |||:  :|  |          :|::| |:||     ||
Db     229 ----ATVTLAAVMGAFIICWFPYFTAFVYRGLRGDDAINEMLEAIVLWLGYANSALNPIL 284

Qy     341 YGCLNRQIRGELSKQFVCFFKPAPEEELRLPSREG---SIEENFLQ 383
           |  |||  |    : | |           || :|     |:  |  |
Db     285 YAALNRDFRTGYQQLFCC----------RLANRNSHKTSLRSNASQ 320


RESULT 10
US-09-473-634-13
; Sequence 13, Application US/09473634
; Patent No. 6440670
; GENERAL INFORMATION:
;  APPLICANT: Heath, Paul Roy
;  APPLICANT:  Orange, Paul Richard
;  APPLICANT:  Pearson, Ronald Carl Alan
;  APPLICANT:  Wright, Simon Ralph
;  TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO THE DETECTION OF VARIATIONS IN
;  TITLE OF INVENTION:  HUMAN HISTAMINE H2 RECEPTORS
;  FILE REFERENCE: 09347/002001
;  CURRENT APPLICATION NUMBER: US/09/473,634
;  CURRENT FILING DATE:  1999-12-28
;  PRIOR APPLICATION NUMBER: US 08/875,540
;  PRIOR FILING DATE: 1998-01-05
;  PRIOR APPLICATION NUMBER: PCT/EP96/00397
;  PRIOR FILING DATE: 1997-01-30
;  PRIOR APPLICATION NUMBER: GB9503866.7
;  PRIOR FILING DATE: 1995-01-30
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-473-634-13

  Query Match             15.0%;  Score 347;  DB 4;  Length 348;
  Best Local Similarity   28.6%;  Pred. No. 9.9e-20;
  Matches   99;  Conservative   57;  Mismatches  132;  Indels   58;  Gaps   10;

Qy      49 MLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFD 106
           : :| |  |||| |   :     ||     |:  | : |||  | ::|      ||::
Db      22 LAVLILITVAGNVVVCLAVGLNRRLRNLTNCFIVSLAITDLLLGLLVLPF-----SAIYQ 76

Qy     107 HAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVL 163
            :    ||:| | :|  | |   : :||::  |:::||  |: |:|| | :|   ||  |
Db      77 LSCKWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVMDPLRYPVLVTPARVAISL 136

Qy     164 VGVWVKALAMASVPV-LGRVSWEEGAPS--VPPGCSLQWSHSAYCQLFVVVFAVLYFLLP 220
           | :|| :: :: : : ||  |  | :      |::| :     ::: :|  :: | ||
Db     137 VLIWVISITLSFLSIHLGWNSRNETSKGNHTTSKCNVQVN-----EVYGLVDGLVTFYLP 191

Qy     221 LLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTF 280
           ||:: : |  :|||||  | :   : :|                    |     ||
Db     192 LLIMCITYYRIFRVARDQAKRIDHISSWK--------------------AATIREHR-- 228

Qy     281 GGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFF 340
               | | | || | |::|| |||:  :|  |          :|::| |:||     ||
Db     229 ----ATVTLAAVMGAFIICWFPYFTAFVYRGLRGDDAINEMLEAIVLWLGYANSALNPIL 284

Qy     341 YGCLNRQIRGELSKQFVCFFKPAPEEELRLPSREG---SIEENFLQ 383
           |  |||  |    : | |           || :|     |:  |  |
Db     285 YAALNRDFRTGYQQLFCC----------RLANRNSHKTSLRSNASQ 320


RESULT 11
US-08-103-170-4
; Sequence 4, Application US/08103170
; Patent No. 5885824
; GENERAL INFORMATION:
;   APPLICANT:  Yamada, Tadataka
;   APPLICANT:  Gantz, Ira
;   TITLE OF INVENTION:  Recombinant Genomic Clones Encoding
;   TITLE OF INVENTION:  Histamine H1, H2, and H3 Receptors, Methods For Production
;   TITLE OF INVENTION:  Thereof, and Proteins Encoded Therefrom
;   NUMBER OF SEQUENCES:  41
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;     ADDRESSEE:  P.C.
;     STREET:  1755 Jefferson Davis Highway, Fourth Floor
;     CITY:  Arlington
;     STATE:  Virginia
;     COUNTRY:  U.S.A.
;     ZIP:  22202
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/103,170
;     FILING DATE:
;     CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 07/633,060
;     FILING DATE:  24-DEC-1990
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Lavalleye, Jean-Paul
;     REGISTRATION NUMBER:  31,451
;     REFERENCE/DOCKET NUMBER:  2363-017-55
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (703)521-4500
;     TELEFAX:  (703)486-2347
;     TELEX:  24885 OPAT UR
;  INFORMATION FOR SEQ ID NO:  4:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-103-170-4

  Query Match             14.9%;  Score 346;  DB 2;  Length 359;
  Best Local Similarity   28.7%;  Pred. No. 1.2e-19;
  Matches  100;  Conservative   55;  Mismatches  141;  Indels   52;  Gaps   10;

Qy      43 SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLS 100
           || |  ::|:    :|||  |   :     ||     |:  | : |||  | ::|
Db      22 SVVLTVLILI---TIAGNVVVCLAVGLNRRLRSLTNCFIVSLAITDLLLGLLVLPF---- 74

Qy     101 SSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLG 157
            || :  :    ||:| | :|  | |   : :||::  |:::||  |  |:|| | :|
Db      75 -SAFYQLSCRWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPV 133

Qy     158 LVASVLVGVWVKALAMASVPV-LGRVSWEEGAP--SVPPGCSLQWSHSAYCQLFVVVFAV 214
            ||  || :|| :: :: : : ||  |  | :      | | :| :      :: :|  :
Db     134 RVAVSLVLIWVISITLSFLSIHLGWNSRNETSSFNHTIPKCKVQVN-----LVYGLVDGL 188

Qy     215 LYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQT 274
           : | ||||:: : |  :|::||    |   : :|
Db     189 VTFYLPLLVMCITYYRIFKIARRIHDQAKHMGSWKAA---------------------- 225

Qy     275 TPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCF 334
             | |  || | | || | |::|| |||:  :|  |          |:|| |:||
Db     226 ----TIGEHKATVTLAAVMGAFIICWFPYFTVFVYRGLKGDDAINEAFEAVVLWLGYANS 281

Qy     335 TSNPFFYGCLNRQIRGELSKQFVCFFKPAPE--EELRLPSREGSIEEN 380
           ||  |  |||  |   : | |  :||   :|. | |    :  |
Db     282 ALNPILYATLNRDFRTAYQQLFRC--RPASHNAQETSLRSNSSQLARN 327


RESULT 12
US-08-103-170-6
; Sequence 6, Application US/08103170
; Patent No. 5885824
;  GENERAL INFORMATION:
;    APPLICANT:  Yamada, Tadataka
;    APPLICANT:  Gantz, Ira
;    TITLE OF INVENTION:  Recombinant Genomic Clones Encoding
;    TITLE OF INVENTION:  Histamine H1, H2, and H3 Receptors, Methods For Production
;    TITLE OF INVENTION:  Thereof, and Proteins Encoded Therefrom
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 Jefferson Davis Highway, Fourth Floor
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/103,170
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/633,060
;      FILING DATE:  24-DEC-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lavalleye, Jean-Paul
;      REGISTRATION NUMBER:  31,451
;      REFERENCE/DOCKET NUMBER:  2363-017-55
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)521-4500
;      TELEFAX:  (703)486-2347
;      TELEX:  24885 OPAT UR
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-103-170-6

  Query Match             14.9%;  Score 346;  DB 2;  Length 359;
  Best Local Similarity   28.7%;  Pred. No. 1.2e-19;
  Matches  100;  Conservative   55;  Mismatches  141;  Indels   52;  Gaps   10;

Qy      43 SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLS 100
           || |  ::|:    :|||  |   :     ||     |:  | : |||  | ::|
Db      22 SVVLTVLILI---TIAGNVVVCLAVGLNRRLRSLTNCFIVSLAITDLLLGLLVLPF---- 74

Qy     101 SSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLG 157
            || :  :    ||:| | :|  | |   : :||::  |:::||  |  |:|| | :|
Db      75 -SAFYQLSCRWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPV 133

Qy     158 LVASVLVGVWVKALAMASVPV-LGRVSWEEGAP--SVPPGCSLQWSHSAYCQLFVVVFAV 214
            ||  || :|| :: :: : : ||  |  | :      | | :| :      :: :|  :
Db     134 RVAVSLVLIWVISITLSFLSIHLGWNSRNETSSFNHTIPKCKVQVN-----LVYGLVDGL 188

Qy     215 LYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQT 274
           : | ||||:: : |  :|::||    |   : :|
Db     189 VTFYLPLLVMCITYYRIFKIARRIHDQAKHMGSWKAA---------------------- 225

Qy     275 TPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCF 334
             | |  || | | || | |::|| |||:  :|  |          |:|| |:||
Db     226 ----TIGEHKATVTLAAVMGAFIICWFPYFTVFVYRGLKGDDAINEAFEAVVLWLGYANS 281

Qy     335 TSNPFFYGCLNRQIRGELSKQFVCFFKPAPE--EELRLPSREGSIEEN 380
           ||  |  |||  |   : | |  :||   :|  | |    :  |
Db     282 ALNPILYATLNRDFRTAYQQLFRC--RPASHNAQETSLRSNSSQLARN 327


RESULT 13
US-08-748-485-5
; Sequence 5, Application US/08748485
; Patent No. 5817480
;  GENERAL INFORMATION:
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Murry, Lynn E.
;    TITLE OF INVENTION:  NOVEL HISTAMINE H2 RECEPTOR
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,485
;      FILING DATE:  Herewith
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0159 US
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  791239
US-08-748-485-5

  Query Match             14.8%;  Score 343;  DB 2;  Length 359;
  Best Local Similarity   27.0%;  Pred. No. 2.1e-19;
  Matches   99;  Conservative   57;  Mismatches  146;  Indels   64;  Gaps   11;

Qy      49 MLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFD 106
           :::| |  |||| |   :    ||     |:  | : |||  | ::|      ||::
Db      25 LIILILVTVAGNVVVCLAVGLNRRLRSLTNCFIVSLAVTDLLLGLLVLPF-----SAIYQ 79

Qy     107 HAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVL 163
             :    | :| | :|  | |    : :|||::  |:::||  |  |:|| | :|   ||  |
Db      80 LSCKWSFSKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPARVAISL 139

Qy     164 VGVWVKALAMASVPVLGRVSWEEGAPSVPPG-----CSLQWSHSAYCQLFVVVFAVLYFL 218
           | :|| :: :: : :   : |     :          | :| :     ::: :|  :: |
Db     140 VFIWVISITLSFLSI--HLGWNSRNETSKDNDTIVKCKVQVN-----EVYGLVDGLVTFY 192

Qy     219 LPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHR 278
           ||||:: : |  :|::||  | :   : :|
Db     193 LPLLIMCITYFRIFKIAREQARRINHIGSWKAA-------------------------- 225

Qy     279 TFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNP 338
           |    || | | || | |::|| |||:  :|  |          | || |:||     ||
Db     226 TIREHKATVTLAAVMGAFIICWFPYFTVFVYRGLKGDDAVNEVFEDVVLWLGYANSALNP 285

Qy     339 FFYGCLNRQIRGELSKQFVCFFKPAPEEELRLP---SREGSIEENFLQFLQGTGCPSESW 395
             | |||  |    : | |          ||    | | |:  |  | |  : |    |
Db     286 ILYAALNRDFRTAYHQLFCC----------RLASHNSHETSLRLNNSQ-LNRSQCQEPRW 334

Qy     396 -VSRPL 400
            :||
Db     335 QEDKPL 340


RESULT 14
US-08-467-568-13
; Sequence 13, Application US/08467568
; Patent No. 5817477
;  GENERAL INFORMATION:
;    APPLICANT:  SOPPET, DANIEL R
;    TITLE OF INVENTION:  ADRENERGIC RECEPTOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan, Cecchi,
;      ADDRESSEE:  Stewart & Olstein
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07068-1739
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/467,568
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferraro, Gregory D
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-324
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-467-568-13

  Query Match             14.7%;  Score 340;  DB 2;  Length 359;
  Best Local Similarity   28.0%;  Pred. No. 3.6e-19;
  Matches   97;  Conservative   57;  Mismatches  134;  Indels   58;  Gaps    9;

Qy      49 MLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFD 106
           : :| |  |||| |   :    ||     |:  | : |||  | ::|      ||::
Db      25 LAVLILITVAGNVVVCLAVGLNRRLRNLTNCFIVSLAITDLLLGLLVLPF-----SAIYQ 79

Qy     107 HAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVL 163
             :    ||:| | :|  | |    : :|||::  |:::||  |: |:|| | :|   ||  |
Db      80 LSCKWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVMDPLRYPVLVTPVRVAISL 139

Qy     164 VGVWVKALAMASVPV-LGRVSWEEGAPS--VPPGCSLQWSHSAYCQLFVVVFAVLYFLLP 220
           | :|| :: :: : : ||  |  | | :        | :| :     ::: :|  :: | ||
Db     140 VLIWVISITLSFLSIHLGWNSRNETSKGNHTTSKCKVQVN-----EVYGLVDGLVTFYLP 194

Qy     221 LLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTF 280
           ||:: : |  :|:|||  | :   : :|                                |
Db     195 LLIMCITYYRIFKVARDQAKRINHISSWKAA--------------------------TI 227

Qy     281 GGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFF 340
             || | | || | |::|| |||:  :|  |          :|::| |:||     ||
Db     228 REHKATVTLAAVMGAFIICWFPYFTAFVYRGLRGDDAINEVLEAIVLWLGYANSALNPIL 287

Qy     341 YGCLNRQIRGELSKQFVCFFKPAPEEELRLPSREG---SIEENFLQ 383
           |  |||  |    : | |          || :|     |:  |  |
Db     288 YAALNRDFRTGYQQLFCC----------RLANRNSHKTSLRSNASQ 323


RESULT 15
US-08-748-485-4
; Sequence 4, Application US/08748485
; Patent No. 5817480
;  GENERAL INFORMATION:
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Murry, Lynn E.
;    TITLE OF INVENTION:  NOVEL HISTAMINE H2 RECEPTOR
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,485
;      FILING DATE:  Herewith
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
```

; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0159 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 184088
US-08-748-485-4

Query Match 14.7%; Score 340; DB 2; Length 359;
Best Local Similarity 28.0%; Pred. No. 3.6e-19;
Matches 97; Conservative 57; Mismatches 134; Indels 58; Gaps 9;

```
Qy      49 MLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFD 106
           : :| |  |||| |   :     ||     |:  | : |||  | ::|      ||::
Db      25 LAVLILITVAGNVVVCLAVGLNRRLRNLTNCFIVSLAITDLLLGLLVLPF-----SAIYQ 79

Qy     107 HAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVL 163
            :    ||:| | :|  | |   : :||::  |:::||  |: |:|| | :|   ||  |
Db      80 LSCKWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVMDPLRYPVLVTPVRVAISL 139

Qy     164 VGVWVKALAMASVPV-LGRVSWEEGAPS--VPPGCSLQWSHSAYCQLFVVVFAVLYFLLP 220
           | :|| :: :: : : ||  |  | :        | :| :     ::: :|  :: | ||
Db     140 VLIWVISITLSFLSIHLGWNSRNETSKGNHTTSKCKVQVN-----EVYGLVDGLVTFYLP 194

Qy     221 LLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTF 280
           ||:: : |  :|:|||  | :   : :|                                |
Db     195 LLIMCITYYRIFKVARDQAKRINHISSWKAA--------------------------TI 227

Qy     281 GGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFF 340
              || | | || | |::|| |||:  :|  |         :|:::| |:||     ||
Db     228 REHKATVTLAAVMGAFIICWFPYFTAFVYRGLRGDDAINEVLEAIVLWLGYANSALNPIL 287

Qy     341 YGCLNRQIRGELSKQFVCFFKPAPEEELRLPSREG---SIEENFLQ 383
           |  |||  |.   : | |          || :|     |:  |  |
Db     288 YAALNRDFRTGYQQLFCC----------RLANRNSHKTSLRSNASQ 323
```

Search completed: December 10, 2002, 09:59:16
Job time : 20 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: December 10, 2002, 09:53:25 ; Search time 14 Seconds
(without alignments)
1336.130 Million cell updates/sec

Title: US-09-838-028-2
Perfect score: 2318
Sequence: 1 MESSPIPQSSGNSSTLGRVP.....SDIIMSDSYLRPAASPRLES 451

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2106	90.9	417	1	GP61_HUMAN	Q9bzj8 homo sapien
2	823	35.5	428	1	GPRX_ORYLA	Q91178 oryzias lat
3	438.5	18.9	368	1	GP62_HUMAN	Q9bzj7 homo sapien
4	351	15.1	358	1	HH2R_RAT	P25102 rattus norv
5	349.5	15.1	440	1	5H6_HUMAN	P50406 homo sapien
6	349	15.1	359	1	HH2R_CANFA	P17124 canis famil
7	345	14.9	358	1	HH2R_MOUSE	P97292 mus musculu
8	345	14.9	576	1	A1AD_RABIT	O02666 oryctolagus
9	344	14.8	440	1	5H6_MOUSE	Q9r1c8 mus musculu
10	343	14.8	359	1	HH2R_CAVPO	P47747 cavia porce
11	340	14.7	359	1	HH2R_HUMAN	P25021 homo sapien
12	338.5	14.6	436	1	5H6_RAT	P31388 rattus norv
13	337	14.5	400	1	B3AR_RAT	P26255 rattus norv
14	335.5	14.5	405	1	B3AR_BOVIN	P46626 bos taurus
15	335.5	14.5	561	1	A1AD_RAT	P23944 rattus norv
16	335	14.5	428	1	B4AR_MELGA	P43141 meleagris g
17	332.5	14.3	572	1	A1AD_HUMAN	P25100 homo sapien
18	326.5	14.1	400	1	B3AR_MOUSE	P25962 mus musculu
19	323.5	14.0	405	1	B3AR_CAPHI	Q9xt57 capra hircu
20	322.5	13.9	418	1	B3AR_MACMU	Q28524 macaca mula
21	321.5	13.9	398	1	B3AR_FELCA	Q9tst4 felis silve
22	321.5	13.9	405	1	B3AR_SHEEP	Q9xt58 ovis aries
23	320.5	13.8	562	1	A1AD_MOUSE	P97714 mus musculu
24	319.5	13.8	408	1	B3AR_HUMAN	P13945 homo sapien
25	317.5	13.7	379	1	GRE2_BALAM	Q93127 balanus amp
26	315	13.6	351	1	B3AR_CAVPO	Q60483 cavia porce
27	312.5	13.5	405	1	B3AR_CANFA	O02662 canis famil
28	307	13.2	458	1	YQNJ_CAEEL	O02213 caenorhabdi
29	306.5	13.2	519	1	A1AB_HUMAN	P35368 homo sapien
30	305	13.2	515	1	A1AB_MESAU	P18841 mesocricetu
31	301	13.0	515	1	A1AB_RAT	P15823 rattus norv
32	300.5	13.0	466	1	A1AA_CAVPO	Q9wu25 cavia porce
33	300	12.9	466	1	A1AA_HUMAN	P35348 homo sapien
34	299.5	12.9	539	1	DOP2_DROME	Q24563 drosophila
35	298.5	12.9	466	1	A1AA_RABIT	O02824 oryctolagus
36	295.5	12.7	514	1	A1AB_MOUSE	P97717 mus musculu
37	293.5	12.7	478	1	OPN4_HUMAN	Q9uhm6 homo sapien
38	292	12.6	466	1	A1AA_RAT	P43140 rattus norv
39	291.5	12.6	388	1	5H4_HUMAN	Q13639 homo sapien
40	291	12.6	470	1	A1AA_ORYLA	Q91175 oryzias lat
41	290	12.5	466	1	A1AA_BOVIN	P18130 bos taurus
42	289.5	12.5	388	1	SSR4_HUMAN	P31391 homo sapien
43	289	12.5	446	1	DADR_MACMU	O77680 macaca mula
44	288	12.4	452	1	A2AA_BOVIN	Q28838 bos taurus
45	286.5	12.4	388	1	5H4_CAVPO	O70528 cavia porce

ALIGNMENTS

RESULT 1
GP61_HUMAN
ID GP61_HUMAN STANDARD; PRT; 417 AA.
AC Q9BZJ8; Q96PR4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable G protein-coupled receptor GPR61 (Biogenic amine receptor-
DE like G-protein-coupled receptor).
GN GPR61 OR BALGR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21105913; PubMed=11165367;
RA Lee D.K., George S.R., Cheng R., Nguyen T., Liu Y., Brown M.,
RA Lynch K.R., O'Dowd B.F.;
RT "Identification of four novel human G protein-coupled receptors
RT expressed in the brain.";
RL Brain Res. Mol. Brain Res. 86:13-22(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=21548835; PubMed=11690637;
RA Cikos S., Gregor P., Koppel J.;
RT "Cloning of a novel biogenic amine receptor-like G protein-coupled
RT receptor expressed in human brain.";
RL Biochim. Biophys. Acta 1521:66-72(2001).
CC -!- FUNCTION: Orphan receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in brain; detected in frontal and
CC temporal lobes, occipital pole, amygdala and hippocampus. Also
CC expressed in testis. Low expression in many other tissues.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 415.
CC ---

CC ---
DR EMBL; AF317652; AAK12637.1; -.
DR EMBL; AF258342; AAK97826.1; ALT_FRAME.
DR Genew; HGNC:13300; GPR61.
DR MIM; 606916; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.

```
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN       1     44       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    45     67       1 (POTENTIAL).
FT   DOMAIN       68     75       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     76     98       2 (POTENTIAL).
FT   DOMAIN       99    112       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    113    135       3 (POTENTIAL).
FT   DOMAIN      136    155       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    156    178       4 (POTENTIAL).
FT   DOMAIN      179    206       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    207    229       5 (POTENTIAL).
FT   DOMAIN      230    287       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    288    310       6 (POTENTIAL).
FT   DOMAIN      311    324       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    325    344       7 (POTENTIAL).
FT   DOMAIN      345    417       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     12     12       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     91     91       L -> P (IN REF. 2).
FT   CONFLICT    102    102       P -> S (IN REF. 2).
FT   CONFLICT    113    113       V -> M (IN REF. 2).
FT   CONFLICT    194    194       H -> G (IN REF. 2).
FT   CONFLICT    226    226       L -> V (IN REF. 2).
FT   CONFLICT    241    242       PD -> QH (IN REF. 2).
SQ   SEQUENCE   417 AA;  45606 MW;  3475BD4F1A714FF2 CRC64;

  Query Match             90.9%;  Score 2106;  DB 1;  Length 417;
  Best Local Similarity   98.8%;  Pred. No. 1.2e-123;
  Matches  409;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGN 60

Qy      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db      61 AAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSPALFDHALFGEVACRLYLF 120

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAM 240
           |||||||||||| ||||||||||||||||||||||||||:||||||||||||||||||||
Db     181 RVSWEEGAPSVPPHCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILLVYCSMFRVARVAAM 240

Qy     241 QHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 PDGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCW 300

Qy     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 LPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFF 360

Qy     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRI 414
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRI 414


RESULT 2
GPRX_ORYLA
ID   GPRX_ORYLA     STANDARD;      PRT;   428 AA.
AC   Q91178;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Probable G protein-coupled receptor (Fragment).
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95275927; PubMed=7756357;
RA   Yasuoka A., Abe K., Saigo K., Arai S., Emori Y.;
RT   "Molecular cloning of a fish gene encoding a novel seven-transmembrane
RT   receptor related distantly to catecholamine, histamine, and serotonin
RT   receptors.";
RL   Biochim. Biophys. Acta 1235:467-469(1995).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D43633; BAA07741.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN       1     46       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    47     67       1 (POTENTIAL).
FT   DOMAIN       68     93       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     94    114       2 (POTENTIAL).
FT   DOMAIN      115    120       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    121    141       3 (POTENTIAL).
FT   DOMAIN      142    162       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    163    183       4 (POTENTIAL).
FT   DOMAIN      184    210       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    211    231       5 (POTENTIAL).
FT   DOMAIN      232    293       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    294    314       6 (POTENTIAL).
FT   DOMAIN      315   >428       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      395    398       POLY-SER.
FT   CARBOHYD     18     18       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER     428    428
SQ   SEQUENCE   428 AA;  47259 MW;  B749E8765B58A38F CRC64;

  Query Match             35.5%;  Score 823;  DB 1;  Length 428;
  Best Local Similarity   42.3%;  Pred. No. 2.6e-44;
  Matches  184;  Conservative   81;  Mismatches  142;  Indels   28;  Gaps   11;

Qy       2 ESSPI---PQSSGNSST--LGRVPQTPGPSTASGVPEVGL----RDVASESVALFFMLLL 52
           ::|||:     |  | ||   |  |  |        |:||:    :    :   || |: |
Db       5 KTSPMITSDHSISNFSTGLFGPHPTVP--------PDVGVVTSSQSQMKDLFGLFCMVTL 56

Qy      53 DLTAVAGNAAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGE 112
           :| |:  |  ||  ||: | |:|| || ||| ||:| |: |||| ::|||  |   :|
Db      57 NLIALLANTGVMVAIARAPHLKKFAFVCHLCAVDVLCAILLMPLGIISSSPFFGTVVFTI 116

Qy     113 VACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALA 172
           : |::|:||:|  : |:||:::||:||||:|:||||||||:||: ||  |:: :| |:|
Db     117 LECQVYIFLNVFLIWLSILTITAISVERYFYIVHPMRYEVKMTINLVIGVMLLIWFKSLL 176

Qy     173 MASVPVLGRVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMF 232
           :| | : |   : . :      |||  |||    :| |:| |: || |:::|  || :::
Db     177 LALVTLFGWPPYGHQSSIAASHCSLHASHSRLRGVFAVLFCVICFLAPVVVIFSVYSAVY 236

Qy     233 RVARVAAMQHGP-LPTWMET--PRQRSESLSSRSTMVTSSGAPQ-TTPHRTFGGGKAAVV 288
           :||| ||:|  | :||| :    : ||:|::|::|::|:   ||  :| | | |||||:
Db     237 KVARSAALQQVPAVPTWADASPAKDRSDSINSQTTIITTRTLPQRLSPERAFSGGKAALT 296

Qy     289 LLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQI 348
           |  : ||||:||||:| ||| ::|:    | | :|  | |: |  |  || ||| |||||
Db     297 LAFIVGQFLVCWLPFFIFHLQMSLTGSMKSPGDLEEAVNWLAYSSFAVNPSFYGLLNRQI 356

Qy     349 RGELSK-QFVCFFKPAPEEELRLPSREGSIEENFLQFLQGTGCPSESWVSRPLPSPK-QE 406
           | || | :  |  :|    |:   | ||| :||||||:| |   ||:  |    :|:  |
Db     357 RDELVKFRRCCVTQPV---EIGPSSLEGSFQENFLQFIQRTSSSSETHPSFANSNPRNME 413
```

```
Qy     407 PPAVDFRIPGQIAEE 421
           |   :|||||  ||
Db     414 NQA--HKIPGQIPEE 426


RESULT 3
GP62_HUMAN
ID   GP62_HUMAN     STANDARD;      PRT;   368 AA.
AC   Q9BZJ7;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Probable G protein-coupled receptor GPR62.
GN   GPR62.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21105913; PubMed=11165367;
RA   Lee D.K., George S.R., Cheng R., Nguyen T., Liu Y., Brown M.,
RA   Lynch K.R., O'Dowd B.F.;
RT   "Identification of four novel human G protein-coupled receptors
RT   expressed in the brain.";
RL   Brain Res. Mol. Brain Res. 86:13-22(2001).
CC   -!- FUNCTION: Orphan receptor.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in brain; detected in the basal
CC       forebrain, frontal cortex, caudate, putamen, thalamus and
CC       hippocampus.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF317653; AAK12638.1; -.
DR   Genew; HGNC:13301; GPR62.
DR   MIM; 606917; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     18       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     19     39       1 (POTENTIAL).
FT   DOMAIN       40     53       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     54     74       2 (POTENTIAL).
FT   DOMAIN       75     91       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     92    112       3 (POTENTIAL).
FT   DOMAIN      113    129       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    130    150       4 (POTENTIAL).
FT   DOMAIN      151    177       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    178    198       5 (POTENTIAL).
FT   DOMAIN      199    239       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    240    260       6 (POTENTIAL).
FT   DOMAIN      261    272       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    273    293       7 (POTENTIAL).
FT   DOMAIN      294    368       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      239    285       ALA-RICH.
FT   CARBOHYD      3      3       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   368 AA;  37628 MW;  9CFFF95298D12C75 CRC64;

  Query Match             18.9%;  Score 438.5;  DB 1;  Length 368;
  Best Local Similarity   31.8%;  Pred. No. 1.4e-20;
  Matches  127;  Conservative   55;  Mismatches  146;  Indels   71;  Gaps   14;

Qy      40 ASE---SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFVFVFHLCLVDLLAALTLMPL 96
           |||   |: |    :::: |: || |:: |: :|| ||  ::: |||:|||||| ::|||
Db       9 ASEVAGSLGLILAAVVEVGALLGNGALLVVVLRTPGLRDALYLAHLCVVDLLAAASIMPL 68

Qy      97 AMLSSSAL-FDHALFGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMT 155
           :|::          |   ||  |||   :    | |:|: : ||  :|||:|   |
Db      69 GLLAAPPPGLGRVRLGPAPCRAARFLSAALLPACTLGVAALGLARYRLIVHPLRPGSRPP 128

Qy     156 LGLVASVLVGVWVKALAMASVPVLGRVSWEEGAPSVPPGCSLQWSHSAYCQL-------F 208
           |   ||  ||  |  : :: :|       | |  ||         | | :       |
Db     129 PVL---VLTAVWAAAGLLGALSLL-------GPPPAPPPA------PARCSVLAGGLGPF 172

Qy     209 VVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTS 268
           ::|:| | || ||:|  |  :| ||| ||::  | |      | ||:|| || :::
Db     173 RPLWALLAFALPALLLLGAYGGIFVVARRAALR-PPRPA--RGSRLRSDSLDSRLSIL-- 227

Qy     269 SGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTW 328
           |   |     |||||:      |||  |||||       | |      : |: |||
Db     228 ---PPLRPR--LPGGKAALAPALAVGQFAACWLPY-----GCACLAPAARAAEAEAAVTW 277

Qy     329 IGYFCFTSNPFFYGCLNRQIR---GELSKQFVCFFKPAPEEELRLPSREGSIEENFLQFL 385
           : |  | ::|| || | | :|   | ||:: :    | |
Db     278 VAYSAFAAHPFLYGLLQRPVRLALGRLSRRAL----PGPVR------------------- 314

Qy     386 QGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAEETSE 424
           |  ::|  | |    | ||      | :  |:| |
Db     315 ---ACTPQAWHPRALLQCLQRPPEGPAVGPSEAPEQTPE 350


RESULT 4
HH2R_RAT
ID   HH2R_RAT       STANDARD;      PRT;   358 AA.
AC   P25102;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Histamine H2 receptor (H2R) (Gastric receptor I).
GN   HRH2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92028890; PubMed=1930188;
RA   Ruat M., Traiffort E., Arrang J.-M., Leurs R., Schwartz J.-C.;
RT   "Cloning and tissue expression of a rat histamine H2-receptor gene.";
RL   Biochem. Biophys. Res. Commun. 179:1470-1478(1991).
CC   -!- FUNCTION: THE H2 SUBCLASS OF HISTAMINE RECEPTORS MEDIATES GASTRIC
CC       ACID SECRETION. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G
CC       PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; S57565; AAB19935.1; -.
DR   PIR; JQ1278; JQ1278.
DR   HSSP; P29274; 1MMH.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
```

```
FT   DOMAIN        1     22       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     23     44       1 (POTENTIAL).
FT   DOMAIN       45     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     81       2 (POTENTIAL).
FT   DOMAIN       82     92       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     93    114       3 (POTENTIAL).
FT   DOMAIN      115    134       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    135    159       4 (POTENTIAL).
FT   DOMAIN      160    179       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    180    203       5 (POTENTIAL).
FT   DOMAIN      204    233       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    234    257       6 (POTENTIAL).
FT   DOMAIN      258    266       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    267    288       7 (POTENTIAL).
FT   DOMAIN      289    358       CYTOPLASMIC (POTENTIAL).
FT   SITE         98     98       ESSENTIAL FOR HISTAMINE BINDING (BY
FT                                 SIMILARITY).
FT   SITE        185    185       ESSENTIAL FOR TIOTIDINE BINDING AND
FT                                 IMPLICATED IN HISTAMINE BINDING (BY
FT                                 SIMILARITY).
FT   SITE        189    189       IMPLICATED IN HISTAMINE BINDING (BY
FT                                 SIMILARITY).
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     91    173       BY SIMILARITY.
SQ   SEQUENCE   358 AA;  40253 MW;  4889F69B7B5D5DDC CRC64;

  Query Match             15.1%;  Score 351;  DB 1;  Length 358;
  Best Local Similarity   27.5%;  Pred. No. 3.5e-15;
  Matches  106;  Conservative   58;  Mismatches  150;  Indels   72;  Gaps   12;

Qy      43 SVALFFMLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLS 100
           || |  ::|:    :|||  |   ::    ||     |:  |   |||  | ::|
Db      22 SVVLTTLILI---TIAGNVVVCLAVSLNRRLRSLTNCFIVSLAATDLLLGLLVLPF---- 74

Qy     101 SSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLG 157
           ||::  :    || | | :|  | |   : :||::  |:::||  |  |:|| | :|
Db      75 -SAIYQLSFTWSFGHVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLVTPV 133

Qy     158 LVASVLVGVWVKALAMASVPVLGRVSWEE------GAPSVPPGCSLQWSHSAYCQLFVVV 211
           ||  || :|| :: :: : :   : |       |  :    | :| :     ::: :|
Db     134 RVAISLVFIWVISITLSFLSI--HLGWNSRNGTRGGNDTFK--CKVQVN-----EVYGLV 184

Qy     212 FAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGA 271
             :: | ||||:: | |  :|::||  | :   : :|
Db     185 DGLVTFYLPLLIMCVTYYRIFKIAREQAKRINHISSWKAA-------------------- 224

Qy     272 PQTTPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGY 331
                 |    || | | || | |::|| |||:  :|  |          || :| |:||
Db     225 -------TIREHKATVTLAAVMGAFIICWFPYFTAFVYRGLRGDDAINEAVEGIVLWLGY 277

Qy     332 FCFTSNPFFYGCLNRQIRGELSKQFVCFFKP--APEEELRL--------PSREGSIEENF 381
                |  |  |||  |    : | | |    : :  |||         ||||  :|    .
Db     278 ANSALNPILYAALNRDFRTAYQQLFHCKFASHNSHKTSLRLNNSLLPRSQSREGRWQEEK 337

Qy     382 LQFLQGTGCPSESWVSRPLPSPKQEP 407
             ||      |   |  |:  |
Db     338 PLKLQ-------VWSGTELTHPQGNP 356


RESULT 5
5H6_HUMAN
ID   5H6_HUMAN      STANDARD;      PRT;   440 AA.
AC   P50406; Q13640;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   5-hydroxytryptamine 6 receptor (5-HT-6) (Serotonin receptor).
GN   HTR6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Striatum;
RX   MEDLINE=96102917; PubMed=8522988;
RA   Kohen R., Metcalf M.A., Khan N., Druck T., Huebner K., Lachowicz J.E.,
RA   Sibley D.R., Roth B., Hamblin M.W.;
RT   "Cloning, characterization, and chromosomal localization of a human
RT   5-HT6 serotonin receptor.";
RL   J. Neurochem. 66:47-56(1996).
RN   [2]
RP   SEQUENCE OF 215-280 FROM N.A.
RC   TISSUE=Striatum;
RX   MEDLINE=95385798; PubMed=7656980;
RA   Ullmer C., Schmuck K., Kalkman H.O., Lubbert H.;
RT   "Expression of serotonin receptor mRNAs in blood vessels.";
RL   FEBS Lett. 370:215-221(1995).
CC   -!- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR
CC       5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTIONS
CC       AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF
CC       THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT STIMULATES ADENYLATE
CC       CYCLASE. IT HAS A HIGH AFFINITY FOR TRICYCLIC PSYCHOTROPIC DRUGS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN SEVERAL HUMAN BRAIN REGIONS, MOST
CC       PROMINENTLY IN THE CAUDATE NUCLEUS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L41147; AAA92622.1; -.
DR   EMBL; Z49119; CAA88929.1; -.
DR   Genew; HGNC:5301; HTR6.
DR   MIM; 601109; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family.
FT   DOMAIN        1     34       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     35     57       1 (POTENTIAL).
FT   DOMAIN       58     64       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     65     85       2 (POTENTIAL).
FT   DOMAIN       86    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    122       3 (POTENTIAL).
FT   DOMAIN      123    144       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    145    166       4 (POTENTIAL).
FT   DOMAIN      167    184       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    185    208       5 (POTENTIAL).
FT   DOMAIN      209    265       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    266    290       6 (POTENTIAL).
FT   DOMAIN      291    295       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    296    320       7 (POTENTIAL).
FT   DOMAIN      321    440       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     99    180       BY SIMILARITY.
FT   CONFLICT    247    247       V -> M (IN REF. 2).
SQ   SEQUENCE   440 AA;  46954 MW;  C888F47650C1D2EF CRC64;

  Query Match             15.1%;  Score 349.5;  DB 1;  Length 440;
  Best Local Similarity   30.7%;  Pred. No. 5.2e-15;
  Matches  127;  Conservative   59;  Mismatches  171;  Indels   57;  Gaps   18;

Qy      19 VPQTPGPSTASGVPEVGLRDVASES----VALFFMLLLDLTAVAGNAAVMAVIAKTPALR 74
           ||: ||| ||:  |  |    ::      ||    ::: ||| | |: ::|:|   ||||
Db       2 VPE-PGP-TANSTPAWGAGPPSAPGGSGWVAAALCVVIALTA-AANSLLIALICTQPALR 58

Qy      75 KF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILS 132
             |:  |   ||:  | :|| |||  :||:   :     | |:    |   | :||:
Db      59 NTSNFFLVSLFTSDLMVGLVVMPPAML--NALYGRWVLARGLCLLWTAFDVMCCSASILN 116
```

```
Qy     133 VSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLGRVSWEE---GAP 189
           :  |:::||  :: |:||::|||     ::::| |  |   : :|:|  : | |     |
Db     117 LCLISLDRYLLILSPLRYKLRMTPLRALALVLGAWSLAALASFLPLL--LGWHELGHARP 174

Qy     190 SVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWM 249
            ||  | |  |       ||:| : | | ||   |   || :   ||  |:|   | | |
Db     175 PVPGQCRLLASLP-----FVLVASGLTFFLPSGAICFTYCRILLAARKQAVQVASLTTGM 229

Qy     250 ----------ETPRQRSESLSSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLL 298
                    |||   ||  ||       |              ||:: |  : | | :
Db     230 ASQASETLQVPRTPRPGVESADSRRLATKHSRKAL----------KASLTLGILLGMFFV 279

Qy     299 CWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVC 358
           |||:|  ::  |:    || |  : |:||:||   | ||  |    |  :  | :
Db     280 TWLPFFVANIVQAV-CDCISPGLFD-VLTWLGYCNSTMNPIIYPLFMRDFKRALGR---- 333

Qy     359 FFKPAPEEELRLP-SREGSIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVD 411
           | | |    | |  |: |:    |:       |  | : : || |   ||  |
Db     334 -FLPCP----RCPRERQASLASPSLRTSHSGPRPGLS-LQQVLPLPL--PPDSD 379


RESULT 6
HH2R_CANFA
ID   HH2R_CANFA     STANDARD;      PRT;   359 AA.
AC   P17124;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Histamine H2 receptor (H2R) (Gastric receptor I).
GN   HRH2.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91110543; PubMed=1703298;
RA   Gantz I., Schaeffer M., Delvalle J., Logsdon C., Campbell V.,
RA   Uhler M., Yamada T.;
RT   "Molecular cloning of a gene encoding the histamine H2 receptor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:429-433(1991).
RN   [2]
RP   MUTAGENESIS OF HISTAMINE-BINDING RESIDUES.
RX   MEDLINE=93015991; PubMed=1356984;
RA   Gantz I., DelValle J., Wang L.-D., Tashiro T., Munzert G., Guo Y.-J.,
RA   Konda Y., Yamada T.;
RT   "Molecular basis for the interaction of histamine with the histamine
RT   H2 receptor.";
RL   J. Biol. Chem. 267:20840-20843(1992).
CC   -!- FUNCTION: THE H2 SUBCLASS OF HISTAMINE RECEPTORS MEDIATES GASTRIC
CC       ACID SECRETION. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G
CC       PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: GASTRIC FUNDUS AND, TO A LESSER EXTENT, IN
CC       BRAIN.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M32701; AAA85637.1; -.
DR   PIR; A39008; A39008.
DR   HSSP; P29274; 1MMH.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     22       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     23     44       1 (POTENTIAL).
FT   DOMAIN       45     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     81       2 (POTENTIAL).
FT   DOMAIN       82     92       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     93    114       3 (POTENTIAL).
FT   DOMAIN      115    134       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    135    159       4 (POTENTIAL).
FT   DOMAIN      160    180       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    181    204       5 (POTENTIAL).
FT   DOMAIN      205    234       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    235    258       6 (POTENTIAL).
FT   DOMAIN      259    267       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    268    289       7 (POTENTIAL).
FT   DOMAIN      290    359       CYTOPLASMIC (POTENTIAL).
FT   SITE         98     98       ESSENTIAL FOR HISTAMINE BINDING.
FT   SITE        186    186       ESSENTIAL FOR TIOTIDINE BINDING AND
FT                                H2 SELECTIVITY.
FT   SITE        190    190       IMPLICATED IN HISTAMINE BINDING.
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     91    174       BY SIMILARITY.
FT   MUTAGEN      98     98       D->N: ABOLISHES HISTAMINE BINDING.
FT   MUTAGEN     186    186       D->A,N: ABOLISHES TIOTIDINE, BUT NOT
FT                                CIMETIDINE BINDING.
FT   MUTAGEN     190    190       T->A,C: DIMINISHED HISTAMINE-STIMULATED
FT                                ACTIVITY.
SQ   SEQUENCE   359 AA;  40201 MW;  37F36412DF5BE805 CRC64;

  Query Match             15.1%;  Score 349;  DB 1;  Length 359;
  Best Local Similarity   28.7%;  Pred. No. 4.6e-15;
  Matches  100;  Conservative   56;  Mismatches  140;  Indels   52;  Gaps   10;

Qy      43 SVALFPMLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLS 100
           || |  ::|:    :|||  |   :     ||     |:  | : |||  | ::|
Db      22 SVVLTVLILI---TIAGNVVVCLAVGLNRRLRSLTNCFIVSLAITDLLLGLLVLPF---- 74

Qy     101 SSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLG 157
            || :  :    ||:| | :|  | |   : :||::  |:::||  |  |:|| | :|
Db      75 -SAFYQLSCRWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPV 133

Qy     158 LVASVLVGVWVKALAMASVPV-LGRVSWEEGAP--SVPPGCSLQWSHSAYCQLFVVVFAV 214
           ||  || :|| :: :: : : ||  |  | :      | | :| :      :: :|  :
Db     134 RVAVSLVLIWVISITLSFLSIHLGWNSRNETSSFNHTIPKCKVQVN-----LVYGLVDGL 188

Qy     215 LYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQT 274
           : | ||||:: : |  :|::||  | :   : :|
Db     189 VTFYLPLLVMCITYYRIFKIARDQAKRIHHMGSWKAA---------------------- 225

Qy     275 TPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCF 334
               | |  || | | '|| | |::|| |||:  :|  |          |:|| |:||
Db     226 ----TIGEHKATVTLAAVMGAFIICWFPYFTVFVYRGLKGDDAINEAFEAVVLWLGYANS 281

Qy     335 TSNPFFYGCLNRQIRGELSKQFVCFFKPAPE--EELRLPSREGSIEEN 380
           ||  |  |||  |    : | |  :||    :|  | |    :  |
Db     282 ALNPILYATLNRDFRTAYQQLFRC--RPASHNAQETSLRSNSSQLARN 327


RESULT 7
HH2R_MOUSE
ID   HH2R_MOUSE     STANDARD;      PRT;   358 AA.
AC   P97292;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Histamine H2 receptor (H2R) (Gastric receptor I).
GN   HRH2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Ola;
RX   MEDLINE=97092891; PubMed=8938453;
RA   Kobayashi T., Inoue I., Jenkins N.A., Gilbert D.J., Copeland N.G.,
RA   Watanabe T.;
RT   "Cloning, RNA expression, and chromosomal location of a mouse
RT   histamine H2 receptor gene.";
RL   Genomics 37:390-394(1996).
CC   -!- FUNCTION: THE H2 SUBCLASS OF HISTAMINE RECEPTORS MEDIATES GASTRIC
CC       ACID SECRETION. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G
CC       PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D50096; BAA08792.1; -.
DR   MGD; MGI:108482; Hrh2.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN        1     22       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     23     44       1 (POTENTIAL).
FT   DOMAIN       45     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     81       2 (POTENTIAL).
FT   DOMAIN       82     92       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     93    114       3 (POTENTIAL).
FT   DOMAIN      115    134       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    135    159       4 (POTENTIAL).
FT   DOMAIN      160    179       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    180    203       5 (POTENTIAL).
FT   DOMAIN      204    233       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    234    257       6 (POTENTIAL).
FT   DOMAIN      258    266       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    267    288       7 (POTENTIAL).
FT   DOMAIN      289    358       CYTOPLASMIC (POTENTIAL).
FT   SITE         98     98       ESSENTIAL FOR HISTAMINE BINDING (BY
FT                                 SIMILARITY).
FT   SITE        185    185       ESSENTIAL FOR TIOTIDINE BINDING AND
FT                                 IMPLICATED IN HISTAMINE BINDING (BY
FT                                 SIMILARITY).
FT   SITE        189    189       IMPLICATED IN HISTAMINE BINDING (BY
FT                                 SIMILARITY).
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     91    173       BY SIMILARITY.
SQ   SEQUENCE   358 AA;  40379 MW;  D3DBA81D71B6D927 CRC64;

  Query Match             14.9%;  Score 345;  DB 1;  Length 358;
  Best Local Similarity   27.6%;  Pred. No. 8.2e-15;
  Matches  108;  Conservative   58;  Mismatches  152;  Indels   74;  Gaps   12;

Qy      42 ESVALFFMLLLDLT-----AVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLM 94
           :|:||   : : ||      ||||  |   ::    ||     |:  |   |||  | :|
Db      13 DSIALKVTISVVLTTLIFITVAGNVVVCLAVSLNRRLRSLTNCFIVSLAATDLLLGLLVM 72

Qy      95 PLAMLSSSALFDHAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYE 151
           |      ||::  :    ||:| | :|  | |   : :||::  |:::||  |  |:||
Db      73 PF-----SAIYQLSFKWRFGQVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYP 127

Qy     152 VRMTLGLVASVLVGVWVKALAMASVPVLGRVSWEE------GAPSVPPGCSLQWSHSAYC 205
           | :|   ||  || :|| :: :: : :   : |       |  :    | :| :
Db     128 VLVTPVRVAISLVFIWVISITLSFLSI--HLGWNSRNGTRGGNDTFK--CKVQVN----- 178

Qy     206 QLFVVVFAVLYFLLPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTM 265
```

```
           ::: :|  :: | ||||:: | |  :|::||  | :   : :|
Db     179 EVYGLVDGMVTFYLPLLIMCVTYYRIFKIAREQAKRINHISSWKAA---------- 224

Qy     266 VTSSGAPQTTPHRTFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESV 325
                       |    || | | || | |::|| |||:  :|  |         || :
Db     225 ------------TIREHKATVTLAAVMGAFIVCWFPYFTAFVYRGLRGDDPVNEVVEGI 271

Qy     326 VTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFFKP--APEEELRL--------PSREG 375
           | |:||     ||  |  |||  |    : | |      : :  |||         ||||
Db     272 VLWLGYANSALNPILYATLNRDFRMAYQQLFHCKLASHNSHKTSLRLNNSLLSRSQSREG 331

Qy     376 SIEENFLQFLQGTGCPSESWVSRPLPSPKQEP 407
           :|     ||       |   |  |:  |
Db     332 RWQEEKPLKLQ-------VWSGTELTHPQGSP 356


RESULT 8
A1AD_RABIT
ID   A1AD_RABIT     STANDARD;      PRT;   576 AA.
AC   O02666;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Alpha-1D adrenergic receptor (Alpha 1D-adrenoceptor).
GN   ADRA1D.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=97182161; PubMed=9030207;
RA   Suzuki F., Miyamoto S., Takita M., Oshita M., Watanabe Y.,
RA   Kakizuka A., Narumiya S., Taniguchi T., Muramatsu I.;
RT   "Cloning, functional expression and tissue distribution of rabbit
RT   alpha 1d-adrenoceptor.";
RL   Biochim. Biophys. Acta 1323:6-11(1997).
CC   -!- FUNCTION: THIS ALPHA-ADRENERGIC RECEPTOR MEDIATES ITS EFFECT
CC       THROUGH THE INFLUX OF EXTRACELLULAR CALCIUM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U64032; AAB53098.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Phosphorylation; Lipoprotein; Palmitate.
FT   DOMAIN        1    101       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    102    126       1 (POTENTIAL).
FT   DOMAIN      127    138       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    139    164       2 (POTENTIAL).
FT   DOMAIN      165    174       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    175    197       3 (POTENTIAL).
FT   DOMAIN      198    218       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    219    243       4 (POTENTIAL).
FT   DOMAIN      244    256       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    257    280       5 (POTENTIAL).
FT   DOMAIN      281    353       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    354    378       6 (POTENTIAL).
FT   DOMAIN      379    385       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    386    410       7 (POTENTIAL).
```

```
FT   DOMAIN      411    576       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     70     70       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   LIPID       424    424       PALMITATE (POTENTIAL).
FT   DOMAIN       54     59       POLY-GLY.
FT   DOMAIN       78     84       POLY-GLY.
FT   DOMAIN      427    433       POLY-ARG.
SQ   SEQUENCE   576 AA;  60266 MW;  21EDF1025D0144CE CRC64;

  Query Match           14.9%;  Score 345;  DB 1;  Length 576;
  Best Local Similarity  25.0%;  Pred. No. 1.3e-14;
  Matches  128;  Conservative   85;  Mismatches  198;  Indels  102;  Gaps   19;

Qy       4 SPIPQSSG--NSSTLGRVPQTPGPSTASGVPEVGLRDVASESVAL-FFMLLLDLTAVAGN 60
Db      60 SVVGAGSGEDNRSSAGEPGGAGGGGEVNGTAAVGGLVVSAQSVGVGVFLAAFILTAVAGN 119

Qy      61 AAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLY 118
Db     120 LLVILSVACNRHLQTVTNYFIVNLAVADLLLSATVLPFS--ATMEVLGFWAFGRAFCDVW 177

Qy     119 LFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPV 178
Db     178 AAVDVLCCTASILSLCTISVDRYVGVRHSLKYPAIMTERKAAAILALLWAVALVVSMGPL 237

Qy     179 LGRVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFR 233
Db     238 LG---WKE---PVPP-------DERFCGITEEVGYAVFSSLCSFYLPMAVIVVMYCRVYV 284

Qy     234 VAR--VAAMQHGPLPTWMETPRQRSES------LSSRSTMVTSSGAPQT--TPHRTFGGG 283
Db     285 VARSTTRSLEAG-------VKRERGKASEVVLRIHCRGAASGADGAPGTRGAKGHTFRSS 337

Qy     284 ------------KAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGY 331
Db     338 LSVRLLKFSREKKAAKTLAIVVGVFVLCWFPFFFVLPLGSLFPQLKPSEGVFKVIFWLGY 397

Qy     332 FCFTSNPFFYGCLNRQIRGELSKQFVCFFK----------------------PAPEEEL 368
Db     398 FNSCVNPLIYPCSSREFKRAFLRLLRCQCRRRRRRRPLWRVYGHHWRASAGGGPHPDCAL 457

Qy     369 ----RLPSREGSIEENFLQFLQGTGCPSESWVSRP----LPSPKQEPPAV--DFRIPGQI 418
Db     458 SAGAALPGAA----------LALTAAPAPSSAAAPEGQAAGAGRRKPPCAFREWRLLGPL 507

Qy     419 AEETSEFLEQQLTSDIIMSDSYLRPAASPRLES 451
Db     508 RRPTT-----QLRAKVSSLSHKIRAGGAQRAEA 535


RESULT 9
5H6_MOUSE
ID   5H6_MOUSE      STANDARD;      PRT;   440 AA.
AC   Q9R1C8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   5-hydroxytryptamine 6 receptor (5-HT-6) (Serotonin receptor).
GN   HTR6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ; TISSUE=Brain;
RX   MEDLINE=21299097; PubMed=11406289;
RA   Kohen R., Fashingbauer L.A., Heidmann D.E.A., Guthrie C.R.,
RA   Hamblin M.W.;
RT   "Cloning of the mouse 5-HT6 serotonin receptor and mutagenesis
RT   studies of the third cytoplasmic loop.";
RL   Brain Res. Mol. Brain Res. 90:110-117(2001).
CC   -!- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR 5-
CC       HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTION AS
CC       A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF THIS
CC       RECEPTOR IS MEDIATED BY G PROTEINS THAT STIMULATES ADENYLATE
CC       CYCLASE. IT HAS A HIGH AFFINITY FOR TRICYCLIC PSYCHOTROPIC DRUGS
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF134158; AAD46490.1; -.
DR   MGD; MGI:1196627; Htr6.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family.
FT   DOMAIN        1     34       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     35     57       1 (POTENTIAL).
FT   DOMAIN       58     64       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     65     85       2 (POTENTIAL).
FT   DOMAIN       86    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    122       3 (POTENTIAL).
FT   DOMAIN      123    144       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    145    166       4 (POTENTIAL).
FT   DOMAIN      167    184       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    185    208       5 (POTENTIAL).
FT   DOMAIN      209    267       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    268    292       6 (POTENTIAL).
FT   DOMAIN      293    297       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    298    322       7 (POTENTIAL).
FT   DOMAIN      323    440       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     99    180       BY SIMILARITY.
FT   CARBOHYD      9      9       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   440 AA;  46998 MW;  4440CDEBE01FEF0C CRC64;

  Query Match           14.8%;  Score 344;  DB 1;  Length 440;
  Best Local Similarity  28.8%;  Pred. No. 1.1e-14;
  Matches  134;  Conservative   59;  Mismatches  187;  Indels   86;  Gaps   19;

Qy       2 ESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNA 61
Db       4 EPGPV-----NSST---PAWGPGPPPAPG---------GSGWVAAALCVVIVLTA-AANS 45

Qy      62 AVMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYL 119
Db      46 LLIALICTQPALRNTSNFFLVSLFTSDLMVGLVVMPPAML--NALYGRWVLARGLCLLWT 103

Qy     120 FLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVL 179
Db     104 AFDVMCCSASILNLCLISLDRYLLILSPLRYKLRMTAPRALALILGAWSLAALASFLPLL 163

Qy     180 GRVSWEE---GAPSVPPGCSLQWSHSAYCQLFVVVFAVLYFLLPLLLILVVYCSMFRVAR 236
Db     164 --LGWHELGKARTSAPGQCRLLASLP-----YVLVASGVTFFLPSGAICFTYCRILLAAR 216

Qy     237 VAAMQHGPLPTWMETPRQRSESL----SSRSTMVTSSGAPQTTPHRTFGGGKAAVVLLAV 292
Db     217 KQAVQVASLTTGTATAGQALETLQVPRTPRPGMESADSRRLTTKH----SRKALKASLTL 272

Qy     293 G---GQFLLCWLPYFSFHLYVALSAQP----ISTGQVESVVTWIGYFCFTSNPFFYGCLN 345
Db     273 GILLSMFFVTWLPFF-----VASIAQAVCDCISPGLFD-VLTWLGYCNSTMNPIIYPLFM 326

Qy     346 RQIRGELSKQFVCFFKPAPEEE------------------------LRLP----SREG 375
```

```
          |  :  | :   |   | ||                                  | ||    |
Db    327 RDFKRALGRFVPCVHCP-PEHRASPASPSMWTSHSGARPGLSLQQVLPLPLPPNSDSDSA 385

Qy    376 SIEENFLQFLQGTGCPSESWVSRPLPSPKQEPPAVDFRIPGQIAEE 421
          |   : ||       | |:  :|  | | : |  |:| :   :  |
Db    386 SGGTSGLQLTAQLLLPGEA--TRDPPPPTRAPTVVNFFVTDSVEPE 429


RESULT 10
HH2R_CAVPO
ID   HH2R_CAVPO     STANDARD;      PRT;   359 AA.
AC   P47747;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Histamine H2 receptor (H2R) (Gastric receptor I).
GN   HRH2.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Hartley; TISSUE=Liver;
RA   Traiffort E.;
RL   Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE H2 SUBCLASS OF HISTAMINE RECEPTORS MEDIATES GASTRIC
CC       ACID SECRETION. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G
CC       PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U25440; AAA65713.1; -.
DR   HSSP; P29274; 1MMH.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein.
FT   DOMAIN       1     22       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    23     44       1 (POTENTIAL).
FT   DOMAIN      45     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    58     81       2 (POTENTIAL).
FT   DOMAIN      82     92       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    93    114       3 (POTENTIAL).
FT   DOMAIN     115    134       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   135    159       4 (POTENTIAL).
FT   DOMAIN     160    180       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   181    204       5 (POTENTIAL).
FT   DOMAIN     205    234       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   235    258       6 (POTENTIAL).
FT   DOMAIN     259    267       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   268    289       7 (POTENTIAL).
FT   DOMAIN     290    359       CYTOPLASMIC (POTENTIAL).
FT   SITE        98     98       ESSENTIAL FOR HISTAMINE BINDING (BY
FT                               SIMILARITY).
FT   SITE       186    186       ESSENTIAL FOR TIOTIDINE BINDING AND
FT                               IMPLICATED IN HISTAMINE BINDING (BY
FT                               SIMILARITY).
FT   SITE       190    190       IMPLICATED IN HISTAMINE BINDING (BY
FT                               SIMILARITY).
FT   CARBOHYD     4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    91    174       BY SIMILARITY.
SQ   SEQUENCE   359 AA;  40556 MW;  58DB81BD8FC3C0E9 CRC64;
```

```
  Query Match          14.8%;  Score 343;  DB 1;  Length 359;
  Best Local Similarity  27.0%;  Pred. No. 1.1e-14;
  Matches   99;  Conservative   57;  Mismatches  146;  Indels   64;  Gaps   11;

Qy     49 MLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFD 106
          :::| |  |||| |   :    ||     |:  | : |||  | ::|      ||::
Db     25 LIILILVTVAGNVVVCLAVGLNRRLRSLTNCFIVSLAVTDLLLGLLVLPF-----SAIYQ 79

Qy    107 HAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVL 163
           :    | :| | :|  | |   : :||::  |:::||  |  |:|| | :|   || |
Db     80 LSCKWSFSKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPARVAISL 139

Qy    164 VGVWVKALAMASVPVLGRVSWEEGAPSVPPG-----CSLQWSHSAYCQLFVVVFAVLYFL 218
          | :|| :: :: : :   : |     :          | :| :     ::: :|  :: |
Db    140 VFIWVISITLSFLSI--HLGWNSRNETSKDNDTIVKCKVQVN-----EVYGLVDGLVTFY 192

Qy    219 LPLLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHR 278
          ||||:: : |  :|::||  | :   : :|
Db    193 LPLLIMCITYFRIFKIAREQARRINHIGSWKAA-------------------------- 225

Qy    279 TFGGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNP 338
          |    || | | || | |::|| |||:  :|  |           | || |:||     ||
Db    226 TIREHKATVTLAAVMGAFIICWFPYFTVFVYRGLKGDDAVNEVFEDVVLWLGYANSALNP 285

Qy    339 FFYGCLNRQIRGELSKQFVCFFKPAPEEELRLP---SREGSIEENFLQFLQGTGCPSESW 395
            | |||  |    : | |          ||    | | |:  |  | | : |    |
Db    286 ILYAALNRDFRTAYHQLFCC----------RLASHNSHETSLRLNNSQ-LNRSQCQEPRW 334

Qy    396 -VSRPL 400
           :||
Db    335 QEDKPL 340


RESULT 11
HH2R_HUMAN
ID   HH2R_HUMAN     STANDARD;      PRT;   359 AA.
AC   P25021; Q14464;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Histamine H2 receptor (H2R) (Gastric receptor I).
GN   HRH2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91337087; PubMed=1714721;
RA   Gantz I., Munzert G., Tashiro T., Schaeffer M., Wang L.-D.,
RA   DelValle J., Yamada T.;
RT   "Molecular cloning of the human histamine H2 receptor.";
RL   Biochem. Biophys. Res. Commun. 178:1386-1392(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=95275318; PubMed=7755641;
RA   Nishi T., Koike T., Oka T., Maeda M., Futai M.;
RT   "Identification of the promoter region of the human histamine H2-
RT   receptor gene.";
RL   Biochem. Biophys. Res. Commun. 210:616-623(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99297624; PubMed=10371214;
RA   Murakami H., Sun-Wada G., Matsumoto M., Nishi T., Wada Y., Futai M.;
RT   "Human histamine H2 receptor gene: multiple transcription initiation
RT   and tissue-specific expression1.";
RL   FEBS Lett. 451:327-331(1999).
RN   [4]
RP   SEQUENCE OF 4-351 FROM N.A., AND POLYMORPHISM.
RC   TISSUE=Brain;
RX   MEDLINE=96414531; PubMed=8817552;
```

```
RA   Orange P.R., Heath P.R., Wright S.R., Pearson R.C.A.;
RT   "Allelic variations of the human histamine H2 receptor gene.";
RL   NeuroReport 7:1293-1296(1996).
RN   [5]
RP   REVIEW.
RX   MEDLINE=98042107; PubMed=9374694;
RA   DelValle J., Gantz I.;
RT   "Novel insights into histamine H2 receptor biology.";
RL   Am. J. Physiol. 273:G987-G996(1997).
CC   -!- FUNCTION: THE H2 SUBCLASS OF HISTAMINE RECEPTORS MEDIATES GASTRIC
CC       ACID SECRETION. ALSO APPEARS TO REGULATE GASTROINTESTINAL MOTILITY
CC       AND INTESTINAL SECRETION. POSSIBLE ROLE IN REGULATING CELL GROWTH
CC       AND DIFFERENTIATION. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY
CC       G PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE AND, THROUGH A SEPARATE
CC       G PROTEIN-DEPENDENT MECHANISM, THE PHOSPHOINOSITIDE/PROTEIN KINASE
CC       (PKC) SIGNALING PATHWAY (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- MISCELLANEOUS: ANTAGONISTS FOR THIS RECEPTOR HAVE PROVEN TO BE
CC       EFFECTIVE THERAPY FOR ACID PEPTIC DISORDERS OF THE
CC       GASTROINTESTINAL TRACT. CERTAIN ANTAGONISTS ARE USED IN THE
CC       TREATMENT OF NEUROPSYCHIATRIC AND NEUROLOGICAL DISEASES SUCH AS
CC       SCHIZOPHRENIA, ALZHEIMER'S DISEASE AND PARKINSONS'S DISEASE.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M64799; AAA58647.1; -.
DR   EMBL; D49783; BAA08618.1; -.
DR   EMBL; AB023486; BAA84279.1; -.
DR   EMBL; X98133; CAA66832.1; -.
DR   PIR; JH0449; JH0449.
DR   HSSP; P29274; 1MMH.
DR   Genew; HGNC:5183; HRH2.
DR   MIM; 142703; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Polymorphism.
FT   DOMAIN        1     22       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     23     44       1 (POTENTIAL).
FT   DOMAIN       45     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     81       2 (POTENTIAL).
FT   DOMAIN       82     92       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     93    114       3 (POTENTIAL).
FT   DOMAIN      115    134       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    135    159       4 (POTENTIAL).
FT   DOMAIN      160    180       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    181    204       5 (POTENTIAL).
FT   DOMAIN      205    234       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    235    258       6 (POTENTIAL).
FT   DOMAIN      259    267       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    268    289       7 (POTENTIAL).
FT   DOMAIN      290    359       CYTOPLASMIC (POTENTIAL).
FT   SITE         98     98       ESSENTIAL FOR HISTAMINE BINDING (BY
FT                                 SIMILARITY).
FT   SITE        186    186       ESSENTIAL FOR TIOTIDINE BINDING AND
FT                                 IMPLICATED IN H2 SELECTIVITY (BY
FT                                 SIMILARITY).
FT   SITE        190    190       IMPLICATED IN HISTAMINE BINDING (BY
FT                                 SIMILARITY).
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     91    174       BY SIMILARITY.
FT   VARIANT     217    217       N -> D.
FT                                 /FTId=VAR_009958.
FT   VARIANT     231    231       K -> R.
FT                                 /FTId=VAR_009959.
FT   VARIANT     268    268       V -> M.
FT                                 /FTId=VAR_009960.
FT   CONFLICT    133    133       V -> A (IN REF. 4).
FT   CONFLICT    175    175       K -> N (IN REF. 4).
FT   CONFLICT    207    207       K -> R (IN REF. 4).
SQ   SEQUENCE   359 AA;  40098 MW;  9835AE2BA60B9B0F CRC64;

  Query Match             14.7%;  Score 340;  DB 1;  Length 359;
  Best Local Similarity   28.0%;  Pred. No. 1.7e-14;
  Matches   97;  Conservative   57;  Mismatches  134;  Indels   58;  Gaps    9;

Qy      49 MLLLDLTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFD 106
           : :| |  |||| |   :     ||     |:  | : |||  | ::|      ||::
Db      25 LAVLILITVAGNVVVCLAVGLNRRLRNLTNCFIVSLAITDLLLGLLVLPF-----SAIYQ 79

Qy     107 HAL---FGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVL 163
            :    ||:| | :|  | |   : :||::  |:::|| |: |:|| | :|   ||  |
Db      80 LSCKWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVMDPLRYPVLVTPVRVAISL 139

Qy     164 VGVWVKALAMASVPV-LGRVSWEEGAPS--VPPGCSLQWSHSAYCQLFVVVFAVLYFLLP 220
           | :|| :: :: : : || |  | :        | :| :     ::: :|  :: | ||
Db     140 VLIWVISITLSFLSIHLGWNSRNETSKGNHTTSKCKVQVN-----EVYGLVDGLVTFYLP 194

Qy     221 LLLILVVYCSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMVTSSGAPQTTPHRTF 280
           ||:: : |  :|:||| | :   : :|                                |
Db     195 LLIMCITYYRIFKVARDQAKRINHISSWKAA--------------------------TI 227

Qy     281 GGGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFF 340
             || | | || | |::|| |||:  :|  |         :|::| |:||     ||
Db     228 REHKATVTLAAVMGAFIICWFPYFTAFVYRGLRGDDAINEVLEAIVLWLGYANSALNPIL 287

Qy     341 YGCLNRQIRGELSKQFVCFFKPAPEEELRLPSREG---SIEENFLQ 383
           |  |||  |    : | |          || :|     |:  |  |
Db     288 YAALNRDFRTGYQQLFCC----------RLANRNSHKTSLRSNASQ 323


RESULT 12
5H6_RAT
ID   5H6_RAT        STANDARD;      PRT;   436 AA.
AC   P31388;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   5-hydroxytryptamine 6 receptor (5-HT-6) (Serotonin receptor)
DE   (ST-B17).
GN   HTR6.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Striatum;
RX   MEDLINE=93196608; PubMed=7680751;
RA   Monsma F.J. Jr., Shen Y., Ward R.P., Hamblin M.W., Sibley D.R.;
RT   "Cloning and expression of a novel serotonin receptor with high
RT   affinity for tricyclic psychotropic drugs.";
RL   Mol. Pharmacol. 43:320-327(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93277562; PubMed=8389146;
RA   Ruat M., Traiffort E., Arrang J.-M., Tardivel-Lacombe J., Diaz J.,
RA   Leurs R., Schwartz J.-C.;
RT   "A novel rat serotonin (5-HT6) receptor: molecular cloning,
RT   localization and stimulation of cAMP accumulation.";
RL   Biochem. Biophys. Res. Commun. 193:268-276(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar;
RA   Martial R.;
RL   Submitted (XXX-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR
```

```
CC       5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTIONS
CC       AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF
CC       THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT STIMULATES ADENYLATE
CC       CYCLASE. IT HAS A HIGH AFFINITY FOR TRICYCLIC PSYCHOTROPIC DRUGS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: LOCALIZED EXCLUSIVELY IN THE CENTRAL NERVOUS
CC       SYSTEM, PREDOMINANTLY IN THE CORPUS STRIATUM BUT ALSO IN VARIOUS
CC       LIMBIC AND CORTICAL REGIONS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L03202; AAA40618.1; -.
DR   EMBL; S62043; AAB26908.1; -.
DR   EMBL; L19656; AAA40611.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family.
FT   DOMAIN       1     34       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     35    57       1 (POTENTIAL).
FT   DOMAIN       58    64       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     65    85       2 (POTENTIAL).
FT   DOMAIN       86    100      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     101   122      3 (POTENTIAL).
FT   DOMAIN       123   144      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     145   166      4 (POTENTIAL).
FT   DOMAIN       167   184      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     185   208      5 (POTENTIAL).
FT   DOMAIN       209   265      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     266   290      6 (POTENTIAL).
FT   DOMAIN       291   295      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     296   320      7 (POTENTIAL).
FT   DOMAIN       321   436      CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     9     9        N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     99    180      BY SIMILARITY.
FT   CONFLICT     57    57       L -> V (IN REF. 1).
FT   CONFLICT     336   436      PCVHCPPEHRPALPPPPCGPLTAVPDQASACSRCCLCLCRQ
FT                               TQIQTPLQGAPRACSSQPSFCCLERPPGTPRHPPGPPLWST
FT                               SLSQTLWSLRYGRIHSVPP -> HASTVPRSTGQPCLPLHV
FT                               DLSQRCQTRPQLQQVLALPLPPNSDSDSASGGTSGLQLTAQ
FT                               LLLPGEATRDPPPPTRATTVVNFFVTDSVEPEIRPHPLSSP
FT                               VN (IN REF. 1).
SQ   SEQUENCE   436 AA;  46922 MW;  6BC00F6A3CBA5FB4 CRC64;

  Query Match             14.6%;  Score 338.5;  DB 1;  Length 436;
  Best Local Similarity   28.4%;  Pred. No. 2.5e-14;
  Matches  129;  Conservative   58;  Mismatches  171;  Indels   97;  Gaps   20;

Qy       2 ESSPIPQSSGNSSTLGRVPQTPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNA 61
           |  |:     ||||       |||  | |          |  ||    ::: ||| | |:
Db       4 EPGPV-----NSST---PAWGPGPPPAPG---------GSGWVAAALCVVIVLTA-AANS 45

Qy      62 AVMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYL 119
           :: :|   ||||     |: |   ||:  | :|| ||| :||:   :      | |:|
Db      46 LLIVLICTQPALRNTSNFFLVSLFTSDLMVGLVVMPPAML--NALYGRWVLARGLCLLWT 103

Qy     120 FLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVL 179
             |  | :||::  |:::||  :: |:||::|||     ::::| |  |   : :|:|
Db     104 AFDVMCCSASILNLCLISLDRYLLILSPLRYKLRMTAPRALALILGAWSLAALASFLPLL 163

Qy     180 GRVSWEE-GAPSVP-PGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMF 232
             : | | |    | ||          |:|     ||:| : : | ||   |   || :
Db     164 --LGWHELGKARTPAPG---------QCRLLASLPFVLVASGVTFFLPSGAICFTYCRIL 212

Qy     233 RVARVAAMQHGPLPTWM----------ETPRQRSESLSSRSTMVTSSGAPQTTPHRTFG 281
             ||  |:|   | |              |||   ||  ||       |
Db     213 LAARKQAVQVASLTTGTAGQALETLQVPRTPRPGMESADSRRLATKHSRKAL-------- 264

Qy     282 GGKAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFY 341
           ||:: |  : | | : |||:|  ::  |:    || |  : |:|||:||   | ||  |
Db     265 --KASLTLGILLGMFFVTWLPFFVANIAQAV-CDCISPGLFD-VLTWLGYCNSTMNPIIY 320

Qy     342 GCLNRQIRGELSKQFVCFFKPAPEEELRLP----------------------SREGSI 377
             |  :  | :   |   | ||    ||                          |:  |
Db     321 PLFMRDFKRALGRFLPCVHCP-PEHRPALPPPPCGPLTAVPDQASACSRCCLCLCRQTQI 379

Qy     378 EENFLQFLQGT--GC---PSESWVSRPLPSPKQEP 407
           :      |||    |   ||   : ||  :|:  |
Db     380 QTP----LQGAPRACSSQPSFCCLERPPGTPRHPP 410


RESULT 13
B3AR_RAT
ID   B3AR_RAT       STANDARD;      PRT;   400 AA.
AC   P26255;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Beta-3 adrenergic receptor.
GN   ADRB3 OR ADRB3R.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92084710; PubMed=1721063;
RA   Muzzin P., Revelli J.-P., Kuhne F., Gocayne J.D., McCombie W.R.,
RA   Venter J.C., Giacobino J.-P., Fraser C.M.;
RT   "An adipose tissue-specific beta-adrenergic receptor. Molecular
RT   cloning and down-regulation in obesity.";
RL   J. Biol. Chem. 266:24053-24058(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92100057; PubMed=1684635;
RA   Granneman J.G., Lahners K.N., Chaudhry A.;
RT   "Molecular cloning and expression of the rat beta 3-adrenergic
RT   receptor.";
RL   Mol. Pharmacol. 40:895-899(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93178631; PubMed=8382630;
RA   Bensaid M., Kaghad M., Rodriguez M., le Fur G., Caput D.;
RT   "The rat beta 3-adrenergic receptor gene contains an intron.";
RL   FEBS Lett. 318:223-226(1993).
CC   -!- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
CC       INDUCED ACTIVATION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
CC       PROTEINS. BETA-3 IS INVOLVED IN THE REGULATION OF LIPOLYSIS AND
CC       THERMOGENESIS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: WHITE AND BROWN ADIPOSE TISSUES, AND DIGESTIVE
CC       TRACT.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M74716; AAA74470.1; -.
DR   EMBL; S73473; AAB20702.1; -.
DR   EMBL; S56481; AAB25520.1; -.
DR   EMBL; S56152; AAB25521.1; -.
```

```
DR   PIR; A41679; A41679.
DR   PIR; S29808; S29808.
DR   PIR; A53281; A53281.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Phosphorylation; Lipoprotein; Palmitate.
FT   DOMAIN        1     36       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     37     60       1 (POTENTIAL).
FT   DOMAIN       61     69       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     70     88       2 (POTENTIAL).
FT   DOMAIN       89    108       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    109    130       3 (POTENTIAL).
FT   DOMAIN      131    152       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    153    175       4 (POTENTIAL).
FT   DOMAIN      176    200       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    201    222       5 (POTENTIAL).
FT   DOMAIN      223    289       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    290    311       6 (POTENTIAL).
FT   DOMAIN      312    323       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    324    344       7 (POTENTIAL).
FT   DOMAIN      345    400       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    107    186       BY SIMILARITY.
FT   LIPID       358    358       PALMITATE (BY SIMILARITY).
FT   CONFLICT    234    235       LL -> FV (IN REF. 1).
SQ   SEQUENCE   400 AA;  43146 MW;  D588540C4B2CE813 CRC64;

  Query Match             14.5%;  Score 337;  DB 1;  Length 400;
  Best Local Similarity   28.2%;  Pred. No. 2.8e-14;
  Matches  120;  Conservative   65;  Mismatches  188;  Indels   52;  Gaps   17;

Qy       4 SPIPQSSGNSSTLGRVPQ-TPGPSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAA 62
           :| |  :|: :     |   |  :  ||:| |          |     || |  | ||
Db       2 APWPHKNGSLAFWSDAPTLDPSAANTSGLPGV-------PWAAALAGALLALATVGGNLL 54

Qy      63 VMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           |:  ||:|| |:     |||  |   ||:  | :||    :: ||  |   |   | |:
Db      55 VITAIARTPRLQTITNVFVTSLATADLVVGLLVMPPG--ATLALTGHWPLGATGCELWTS 112

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           : |  |: :| :: |: |:||  | :|:||   :|     : :| ||: :  ::  |::
Db     113 VDVLCVTASIETLCALAVDRYLAVTNPLRYGTLVTKRRARAAVVLVWIVSATVSFAPIMS 172

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA 235
           :  |  || :    |    |:  |       : :: : : | ||||:| ||  :| ||
Db     173 Q-WWRVGADAEAQECH---SNPRCCSFASNMPYALLSSSVSFYLPLLVMLFVYARVFVVA 228

Qy     236 ----RVAAMQHGPLPTWMETPR--QRSESLSSRSTMVTSSGAPQ--TTPHR--TFGGGKA 285
               |:   : |  |   |:||   || | ::  |   | | |     | |    |  :|
Db     229 KRQRRLLRRELGRFPP-EESPRSPSRSPSPATVGTPTASDGVPSCGRRPARLLPLGEHRA 287

Qy     286 AVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLN 345
             |  : | | |||||:|  ::  ||    :    |   : |:||     ||  | | :
Db     288 LRTLGLIMGIFSLCWLPFFLANVLRALVGPSLVPSGVFIALNWLGYANSAFNPLIY-CRS 346

Qy     346 RQIRGELSKQFVCFFKPAPEEELR--------LPSREGSIEENFLQFLQGTGCPSESWVS 397
             | :  :: :| :     || |        : ||: |    |       |   |
Db     347 PDFR-DAFRRLLCSYGGRGPEEPRVVTFPASPVASRQNSPLNRF------DGYEGE---- 395

Qy     398 RPLPS 402
           || |:
Db     396 RPFPT 400


RESULT 14
B3AR_BOVIN
ID   B3AR_BOVIN     STANDARD;      PRT;   405 AA.
AC   P46626; Q28045; Q9TS17;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Beta-3 adrenergic receptor.
GN   ADRB3.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brown adipose tissue;
RX   MEDLINE=95324546; PubMed=7601122;
RA   Pietri-Rouxel F., Lenzen G., Kapoor A., Drumare M.F., Archimbault P.,
RA   Strosberg A.D., Manning B.S.J.;
RT   "Molecular cloning and pharmacological characterization of the bovine
RT   beta 3-adrenergic receptor.";
RL   Eur. J. Biochem. 230:350-358(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Simmental;
RX   MEDLINE=20292453; PubMed=10834601;
RA   Forrest R.H., Hickford J.G.H.;
RT   "Rapid communication: nucleotide sequences of the bovine, caprine, and
RT   ovine beta3-adrenergic receptor genes.";
RL   J. Anim. Sci. 78:1397-1398(2000).
RN   [3]
RP   SEQUENCE OF 4-106 FROM N.A.
RC   TISSUE=Muscle;
RA   Stoffel B., Meyer H.H.D.;
RL   Submitted (JAN-1993) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 156-298 FROM N.A.
RC   STRAIN=Friesian; TISSUE=Adipose tissue;
RX   MEDLINE=94107292; PubMed=7904157;
RA   Casteilla L., Muzzin P., Revelli J.-P., Ricquier D., Giacobino J.-P.;
RT   "Expression of beta 1- and beta 3-adrenergic-receptor messages and
RT   adenylate cyclase beta-adrenergic response in bovine perirenal adipose
RT   tissue during its transformation from brown into white fat.";
RL   Biochem. J. 297:93-97(1994).
CC   -!- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
CC       INDUCED ACTIVATION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
CC       PROTEINS. BETA-3 IS INVOLVED IN THE REGULATION OF LIPOLYSIS AND
CC       THERMOGENESIS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X85961; CAA59937.1; -.
DR   EMBL; AF109930; AAD26149.1; -.
DR   EMBL; X67214; CAA47654.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Phosphorylation; Lipoprotein; Palmitate.
FT   DOMAIN        1     36       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     37     63       1 (POTENTIAL).
FT   DOMAIN       64     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     73     91       2 (POTENTIAL).
FT   DOMAIN       92    111       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    112    133       3 (POTENTIAL).
```

```
FT   DOMAIN      134    155       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    156    178       4 (POTENTIAL).
FT   DOMAIN      179    203       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    204    225       5 (POTENTIAL).
FT   DOMAIN      226    292       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    293    314       6 (POTENTIAL).
FT   DOMAIN      315    326       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    327    347       7 (POTENTIAL).
FT   DOMAIN      348    405       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    110    189       BY SIMILARITY.
FT   LIPID       361    361       PALMITATE (BY SIMILARITY).
FT   CONFLICT      6      7       PG -> HE (IN REF. 3).
FT   CONFLICT    156    156       A -> T (IN REF. 4).
SQ   SEQUENCE   405 AA;  42903 MW;  C93667DC1CC286F8 CRC64;

  Query Match             14.5%;  Score 335.5;  DB 1;  Length 405;
  Best Local Similarity   28.1%;  Pred. No. 3.5e-14;
  Matches  121;  Conservative   67;  Mismatches  183;  Indels   59;  Gaps   18;

Qy       7 PQSSGNSSTLGRVPQTP--GPST--ASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAA 62
           |   |||| |   |  |   |:|  |||:| |      : ::|   : |  |  | ||
Db       3 PWPPGNSS-LTPWPDIPTLAPNTANASGLPGVPW----AVALAGALLALAVLATVGGNLL 57

Qy      63 VMAVIAKTPALRKF--VFVFHLCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLF 120
           |:  ||:|| |:    |||  |   ||:  | ::|    :: ||  |   |   | |:
Db      58 VIVAIARTPRLQTMTNVFVTSLATADLVVGLLVVPPG--ATLALTGHWPLGVTGCELWTS 115

Qy     121 LSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKALAMASVPVLG 180
           : |  |: :| :: |: |:||  | :|:||   :|     : :| ||| : |::  |::
Db     116 VDVLCVTASIETLCALAVDRYLAVTNPLRYGALVTKRRALAAVVLVWVVSAAVSFAPIMS 175

Qy     181 RVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILVVYCSMFRVA 235
           : |  || :    |    |:  |       : :: : :| ||||::| ||  :| ||
Db     176 K-WWRIGADAEAQRCH---SNPRCCTFASNMPYALLSSSVSFYLPLLVMLFVYARVFVVA 231

Qy     236 ----RVAAMQHGPLPTWMETPR-QRSESLSSRSTMVTSSGAP-------QTTPHRTFGGG 283
               |:   : |  |      |   || |        : :| |        :  | |
Db     232 TRQLRLLRRELGRFPPEESPPAPSRSGSPGLAGPCASPAGVPSYGRRPARLLPLREH--- 288

Qy     284 KAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGC 343
           :|   |  : | | |||||:|  ::  ||    : :|     : |:||     ||  | |
Db     289 RALRTLGLIMGTFTLCWLPFFVVNVVRALGGPSLVSGPTFLALNWLGYANSAFNPLIY-C 347

Qy     344 LNRQIRGELSKQFVCFFKPAPEEELRL--PSREGSIEENFLQFLQGTGCPSESWVSRPLP 401
           :  |    :: :|  :  ||| |    | |                || :  :   |
Db     348 RSPDFRSAF-RRLLC--RCRPEEHLAAASPPR----------------APSGAPTALTSP 388

Qy     402 SPKQEPPAVD 411
           :  :|| :|
Db     389 AGPMQPPELD 398


RESULT 15
A1AD_RAT
ID   A1AD_RAT       STANDARD;      PRT;   561 AA.
AC   P23944;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Alpha-1D adrenergic receptor (Alpha 1D-adrenoceptor) (Alpha-1A
DE   adrenergic receptor) (RA42).
GN   ADRA1D OR ADRA1A.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain cortex;
RX   MEDLINE=91177889; PubMed=1706716;
RA   Lomasney J.W., Cotecchia S., Lorenz W., Leung W.-Y., Schwinn D.A.,
RA   Yang-Feng T.L., Brownstein M., Lefkowitz R.J., Caron M.G.;
RT   "Molecular cloning and expression of the cDNA for the alpha 1A-
RT   adrenergic receptor. The gene for which is located on human
RT   chromosome 5.";
RL   J. Biol. Chem. 266:6365-6369(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95114877; PubMed=7815325;
RA   Schwinn D.A., Johnston G.I., Page S.O., Mosley M.J., Wilson K.H.,
RA   Worman N.P., Campbell S., Fidock M.D., Furness L.M.,
RA   Parry-Smith D.J., Peter B., Bailey D.S.;
RT   "Cloning and pharmacological characterization of human alpha-1
RT   adrenergic receptors: sequence corrections and direct comparison with
RT   other species homologues.";
RL   J. Pharmacol. Exp. Ther. 272:134-142(1995).
CC   -!- FUNCTION: THIS ALPHA-ADRENERGIC RECEPTOR MEDIATES ITS EFFECT
CC       THROUGH THE INFLUX OF EXTRACELLULAR CALCIUM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: VAS DEFERENS, HIPPOCAMPUS, CEREBRAL CORTEX,
CC       AORTA, BRAIN STEM, HEART AND SPLEEN.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M60654; AAA63477.1; -.
DR   EMBL; L31771; AAB59704.1; -.
DR   PIR; A38731; A38731.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Phosphorylation; Lipoprotein; Palmitate.
FT   DOMAIN        1     90       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     91    115       1 (POTENTIAL).
FT   DOMAIN      116    127       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    128    153       2 (POTENTIAL).
FT   DOMAIN      154    163       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    164    186       3 (POTENTIAL).
FT   DOMAIN      187    207       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    208    232       4 (POTENTIAL).
FT   DOMAIN      233    245       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    246    269       5 (POTENTIAL).
FT   DOMAIN      270    342       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    343    367       6 (POTENTIAL).
FT   DOMAIN      368    374       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    375    399       7 (POTENTIAL).
FT   DOMAIN      400    561       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     60     60       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     76     76       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   LIPID       413    413       PALMITATE (POTENTIAL).
FT   DOMAIN       21     57       GLY-RICH.
FT   DOMAIN      416    421       POLY-ARG.
FT   CONFLICT    424    483       AVYGHHWRASTGDARSDCAPSPRIAPPGAPLALTAHPGAGS
FT                                ADTPETQDSVSSSRKPASA -> SLRPPLASLDRRRAFRLR
FT                                PQPSHRSPRGPSSPHCTPGCGLGRHAGDAGFGLQQSKASLR
FT                                (IN REF. 1).
SQ   SEQUENCE   561 AA;  59354 MW;  B6537DCAD4F7BE27 CRC64;

  Query Match             14.5%;  Score 335.5;  DB 1;  Length 561;
  Best Local Similarity   25.7%;  Pred. No. 4.8e-14;
  Matches  125;  Conservative   81;  Mismatches  195;  Indels   85;  Gaps   19;

Qy       6 IPQSSGNSSTLGR----------VPQTPGPSTASGVPEVGLRDVASESVAL-FFMLLLD 53
           :| ::|  : :|             |        :|   ||   |::: | :  |:
Db      42 VPGATGGGAVVGTGSGEDNQSSTGEPGAAASGEVNGSAAVGGLVVSAQGVGVGVFLAAFI 101
```

```
Qy      54 LTAVAGNAAVMAVIAKTPALRKFV--FVFHLCLVDLLAALTLMPLAMLSSSALFDHALFG 111
           ||||||| |:  :|    |:     |: :| : ||| :  ::| :  ::  :     ||
Db     102 LTAVAGNLLVILSVACNRHLQTVTNYFIVNLAVADLLLSAAVLPFS--ATMEVLGFWAFG 159

Qy     112 EVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPMRYEVRMTLGLVASVLVGVWVKAL 171
             | ::  : |   : :|||:  |:|:||  | | ::|   ||    |::|  :|  ||
Db     160 RTFCDVWAAVDVLCCTASILSLCTISVDRYVGVRHSLKYPAIMTERKAAAILALLWAVAL 219

Qy     172 AMASVPVLGRVSWEEGAPSVPPGCSLQWSHSAYCQL-----FVVVFAVLYFLLPLLLILV 226
            ::  |:||   |:|    |||          :| :     : :  :|  | ||: :|:|
Db     220 VVSVGPLLG---WKE---PVPP-------DERFCGITEEVGYAIFSSVCSFYLPMAVIVV 266

Qy     227 VYCSMFRVAR--VAAMQHGPLPTWMETPRQRSE---SLSSRSTMVTSSGAP--QTTPHRT 279
           :|| :: |||    ::: |        | : ||    :  |    :: | |  |::   |
Db     267 MYCRVYVVARSTTRSLEAG----IKREPGKASEVVLRIHCRGAATSAKGYPGTQSSKGHT 322

Qy     280 FGGG-----------KAAVVLLAVGGQFLLCWLPYFSFHLYVALSAQPISTGQVESVVT 327
                        ||| |  | | |:||| |:|     :|  |   :  |  |:
Db     323 LRSSLSVRLLKFSREKKAAKTLAIVVGVFVLCWFPFFFVLPLGSLFPQLKPSEGVFKVIF 382

Qy     328 WIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFFKPAPEEELRLPSREGSIEENFLQFLQG 387
           |:|||    ||  | | :|:           || |    ||   |        |  : |
Db     383 WLGYFNSCVNPLIYPCSSRE-----------FKRAFLRLLRCQCRR---RRRRLWAVYG 427

Qy     388 TGCPSESWVSRP--LPSPKQEPPAVDFRI---PGQIAEETSEFLEQQLTSDIIMSDSYLR 442
               : :  :|    |||:  ||     :   ||  : :| |      | | : |    :
Db     428 HHWRASTGDARSDCAPSPRIAPPGAPLALTAHPGAGSADTPE------TQDSVSSSR--K 479

Qy     443 PAASPR 448
           ||:: |
Db     480 PASALR 485
```

Search completed: December 10, 2002, 09:57:42
Job time : 17 secs